

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.
OM protein - protein search, using sw model
Run on: February 27, 2006, 09:42:30 ; Search time 184 Seconds
(without alignments)
16.715 Million cell updates/sec

Title: US-10-027-015b-2

Perfect score: 40

Sequence: 1 GGIGDGG 7

Scoring table: BLOSUMP62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

- A: Geneseq_21:*
- 1: geneseqp1980s:*
 - 2: geneseqp1990s:*
 - 3: geneseqp2000s:*
 - 4: geneseqp2001s:*
 - 5: geneseqp2002s:*
 - 6: geneseqp2003as:*
 - 7: geneseqp2003bs:*
 - 8: geneseqp2004s:*
 - 9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	7	ADG43177	Adg43177 IGD-motif
2	40	100.0	212	7 ADG54738	Adg54738 Rat prote
3	40	100.0	449	7 ABO72195	Abp72195 Pseudomon
4	40	100.0	1090	5 ABP35656	Abp35656 Fungal ZB
5	39	97.5	146	7 ABM89751	Abm89751 Rice abio
6	39	97.5	157	4 ABB40107	Abb40107 Peptide #
7	39	97.5	157	4 AAM733743	Aam733743 Peptide #
8	39	97.5	157	4 AAM73550	Aam73550 Human bon
9	39	97.5	157	4 AAM60865	Aam60865 Human bra
10	39	97.5	157	4 ABG55277	Abg55277 Human liv
11	39	97.5	157	5 ABG43413	Abg43413 Human pep
12	39	97.5	233	9 ABM96534	Abm96534 M. xanthu
13	39	97.5	239	4 ADM20064	Adm20064 Protein e
14	39	97.5	250	3 AAB58809	Aab58809 Breat an
15	39	97.5	326	7 ABM86774	Abm86774 Rice abio
16	39	97.5	399	4 ABG28641	Abg28641 Novel hum
17	39	97.5	445	4 ADM19814	Adm19814 Protein e
18	39	97.5	525	5 ABP43927	Abp43927 FLJ20261
19	38	95.0	37	8 ABO55390	Abp55390 Human gen
20	38	95.0	427	4 ABB61028	Abb61028 Drosophil
21	38	95.0	430	7 ABO83915	Abp83915 Pseudomon
22	36	90.0	682	4 AAU02902	Aau02902 Angiotens
23	36	90.0	11096	4 AAE10129	Aae10129 Streptomy
24	35	87.5	19	3 AAY54970	Aay54970 Peptide 1

98 34 85.0 102 6 ABM58882 Propionib
99 34 85.0 104 8 ABO60629 Human gen
100 34 85.0 116 7 ABM89504 Rice ablo

ALIGNMENTS

RESULT 1
ADG43177
ID ADG43177 standard; peptide; 7 AA.
XX
AC ADG43177;
XX
DT 26-FEB-2004 (first entry)
XX
DE IGD-motif peptide #SEQ ID 2.
XX
KW Angiogenic composition; isoleucine-glycine-aspartic acid; IGD;
KW angiogenic growth factor; cell migration; angiogenesis; ischaemia; heart;
KW polyvinyl pyrrolidone.
XX
OS Unidentified.
XX
PN WO2003059436-A2.
XX
PD 24-JUL-2003.
XX
PF 23-DEC-2002; 2002WO-US041484.
XX
PR 21-DEC-2001; 2001US-00027015.
XX
PA (CENT-) CENTERPULSE BIOLOGICS INC.
XX
PI Akella R, Ranieri J;
XX
WPI; 2003-697294/66.
XX
PT Angiogenic composition comprising peptides containing isoleucine-glycine-
PT aspartic acid, that stimulate cell migration, and angiogenic growth
PT factor useful for promoting myocardial or peripheral angiogenesis.
XX
PS Claim 2; SEQ ID NO 2; 22pp; English.
XX
CC The invention relates to an angiogenic composition (I) comprising at
CC least one peptide chosen from a group of peptides containing isoleucine-
CC glycine-aspartic acid (IGD-containing peptide), and at least one
CC angiogenic growth factor other than the at least one peptide, where the
CC peptide stimulates cell migration. Also disclosed is a composition (III)
CC that is active for promoting cell migration and/or angiogenesis under
CC cell growth promoting conditions comprising an IGD-containing peptide,
CC and a matrix material. Compositions of the invention are useful for
CC promoting myocardial angiogenesis which involves administering (I) or
CC (III) intramyocardially to an ischaemic area of the heart of an
CC individual, to enhance vascular endothelial cell migration and/or
CC proliferation. They are also useful for promoting peripheral angiogenesis
CC in an ischaemic region of an organ or tissue fed by a peripheral vessel,
CC by enhancing vascular endothelial cell migration and/or proliferation at
CC the ischaemic region, or for enhancing blood flow to an ischaemic tissue
CC of the body, where the composition stimulates vascular endothelial cell
CC migration and/or proliferation sufficient to restore or increase blood
CC flow to the ischaemic tissue. The physiologically acceptable carrier in
CC the composition is polyvinyl pyrrolidone. The current sequence represents
CC an IGD-motif peptide that may be utilised in an angiogenic composition of
CC the invention.
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 40; DB 7; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 GGI GGG 7
RESULT 2
ADE54738
ID ADE54738 standard; protein; 212 AA.
XX
AC ADE54738;
XX
DT 29-JAN-2004 (first entry)
XX
DE Rat Protein AAB71237, SEQ ID NO 543.
XX
KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
OS Rattus norvegicus.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
WPI; 2003-268312/26.
XX
GENBANK; AAB71237.

New composition comprising two or more isolated polypeptides, useful for
preparing a medicament for treating pain in an animal.
Claim 1; Page; 1017pp; English.
The invention discloses a composition comprising two or more isolated rat
or human polynucleotides or a polynucleotide which represents a fragment,
derivative or allelic variation of the nucleic acid sequence. Also
claimed are a vector comprising the novel polynucleotide, a host cell
comprising the vector, a method for identifying a nucleotide sequence
which is differentially regulated in an animal subjected to pain and a
kit to perform the method, an array, a method for identifying an agent
that increases or decreases the expression of the polynucleotide sequence
that is differentially expressed in neuronal tissue of a first animal
subjected to pain, a method for identifying a compound which regulates
the expression of a polynucleotide sequence which is differentially
expressed in an animal subjected to pain, a method for identifying a
compound that regulates the activity of one or more of the
polynucleotides, a method for producing a pharmaceutical composition, a
method for identifying a compound or small molecule that regulates the
activity in an animal of one or more of the polypeptides given in the
specification, a method for identifying a compound useful in treating
pain and a pharmaceutical composition comprising the one or more
polypeptides or their antibodies. The polynucleotide or the compound that
modulates its activity is useful for preparing a medicament for treating
pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
therapy). The sequence presented is a rat protein (shown in Table 2 of
the specification) which is differentially expressed during pain. Note:
The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic form directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences.

Query Match 100.0%; Score 40; DB 7; Length 212;
Sequence 212 AA;

Best Local Similarity 100.0%; Pred. No. 1.3e+02; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
Db 83 GGIGDGG 89

RESULT 3

ABO72195
ID ABO72195 standard; protein; 449 AA.
XX
AC ABO72195;
XX
DT 29-JUL-2004 (first entry)
XX
DE Pseudomonas aeruginosa polypeptide #4370.
XX
KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
OS Pseudomonas aeruginosa.
XX
PN US6551795-B1.
XX
PD 22-APR-2003.
XX
PF 18-FEB-1999; 99US-00252991.
XX
PR 18-FEB-1998; 98US-0074788P.
PR 27-JUL-1998; 98US-0094190P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX
DR WPI; 2003-615309/58.
DR N-PSDB; ABD05766.
XX

Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide, useful as molecular targets for diagnostics, prophylaxis and treatment of pathological conditions resulting from bacterial infection.

Disclosure; SEQ ID NO 20941; 455pp; English.

The invention relates to Pseudomonas aeruginosa polypeptides and the polynucleotides encoding them. The sequences are useful in diagnosis and therapy of pathological conditions, as molecular targets for diagnostics, prophylaxis and treatment of pathological conditions resulting from a bacterial infection, for evaluating a compound, such as a polypeptide, for the ability to bind a P. aeruginosa nucleic acid, as components of effective antibacterial targets, as targets for antibacterial drugs, including anti-P. aeruginosa drugs, as templates for recombinant production of P. aeruginosa-derived peptides or polypeptides, as target components for diagnosis and/or treatment of P. aeruginosa-caused infection, and in detection of P. aeruginosa sequences or other sequences of Pseudomonas species using biochip technology. Sequences ABO7825-ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format from USPTO at seqdata.uspto.gov/sequence.html

Sequence 449 AA;

Query Match 100.0%; Score 40; DB 7; Length 449;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
Db 329 GGIGDGG 335

RESULT 4

ABP35656

ID ABP35656 standard; protein; 1090 AA.

XX
AC ABP35656;
XX
DT 24-JUL-2002 (first entry)
XX
DE Fungal ZBC protein sequence #82.

XX
KW Secondary metabolite; fungus; ZBC gene; zinc binuclear cluster protein; antibacterial; beta-lactam; anti-hypercholesterolaemic; lovastatin;
KW mevastatin; immunosuppressant; cyclosporin A; ergot alkaloid; ergotamine;
KW angiogenesis inhibitor; ovalicin; glucan synthase inhibitor; gliotoxin;
KW fungal toxin; cell surface receptor; plant growth regulator; pigment;
KW insecticide; antineoplastic.

XX
OS Unidentified.

XX
PN WO200224865-A2.

XX
PD 28-MAR-2002.

XX
PF 19-SEP-2001; 2001WO-US029288.

XX
PR 19-SEP-2000; 2000US-0233564P.

XX
PA (MICR-) MICROBIA INC.

XX
PI Holtzman D, Madden K, Maxon M, Sherman A;

XX
DR WPI; 2002-352005/38.
DR N-PSDB; ABN79845.

XX
PT New method for improving the production of a secondary metabolite e.g. antineoplastic agent, ergot alkaloid from a fungus involves modulation of the expression of at least one zinc binuclear cluster protein gene.

PS Disclosure; SEQ ID NO 176; 49pp + Sequence Listing; English.

XX
CC The invention relates to improving the production of a secondary metabolite by a fungus. This involves modulating the expression of at least one ZBC (zinc binuclear cluster protein) gene in a manner to improve the yield of the secondary metabolite. Methods of the invention may be used for improving the production of the secondary metabolite e.g. antibacterial (such as beta-lactam), an anti-hypercholesterolaemic (such as lovastatin or mevastatin), an immunosuppressant (such as cyclosporin A), an ergot alkaloid (such as ergotamine), an angiogenesis inhibitor (such as ovalicin), a glucan synthase inhibitor, gliotoxin family of compounds, a fungal toxin, a modulator of cell surface receptor signalling, a plant growth regulator, a pigment, an insecticide, or an antineoplastic compound. The method results in a decrease in fermentor run-time, a decrease in the size of the fermentor required for the production of equivalent amounts of the secondary metabolite, or a decrease in the biomass required for the production, which translates into decreased waste that must be handled in downstream processing. The sequences given in records ABP35575-ABP35722 represent ZBC proteins. Note: The sequence data for this patent did not form part of the printed specification, but was obtained directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX
SQ Sequence 1090 AA;

Query Match 100.0%; Score 40; DB 5; Length 1090;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
Db 952 GGIGDGG 958

RESULT 5
ABM89751

ID ABM89751 standard; protein; 146 AA.
 XX
 AC ABM89751;
 XX
 DT 02-JUN-2005 (first entry)
 XX
 DE Rice abiotic stress responsive polypeptide SEQ ID NO:7997.
 XX
 KW abiotic stress tolerance; transgenic plant; plant; cereal; agriculture.
 XX
 OS Oryza sativa.
 XX
 PN WO2003008540-A2.
 XX
 PD 30-JAN-2003.
 XX
 PF 21-JUN-2002; 2002WO-US019668.
 XX
 PR 22-JUN-2001; 2001US-0300112P.
 PR 24-AUG-2001; 2001US-0314662P.
 PR 26-SEP-2001; 2001US-0325277P.
 PR 21-NOV-2001; 2001US-0332132P.
 XX
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX
 PI Krepis J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F;
 PI Moughamer T, Provart N, Ricke D, Zhu T;
 XX
 DR WPI; 2003-248011/24.
 XX
 PT New stress-responsive nucleic acid, useful for altering the
 PT responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold
 PT stress, salt stress or osmotic stress.
 XX
 PS Claim 1; SEQ ID NO 7997; 89pp; English.
 XX
 CC The invention relates to novel abiotic stress responsive polynucleotides
 CC and polypeptides. Also disclosed are vectors, expression cassettes, host
 CC cells, and plants containing such polynucleotides. Also disclosed are
 CC methods for using the polynucleotides and polypeptides to alter the
 CC responsiveness of a plant to abiotic stress. The invention is useful in
 CC agriculture. The nucleic acid is useful for determining whether a test
 CC plant has been exposed to an abiotic stress condition. It is also useful
 CC for selecting an agent that alters abiotic stress regulated
 CC polynucleotide expression in a plant cell, and to identify a homolog or
 CC ortholog to an abiotic stress responsive polynucleotide. The nucleic acid
 CC molecule and the polypeptide encoded by it are useful in altering the
 CC responsiveness of a plant to an abiotic stress, such as cold stress, salt
 CC stress, osmotic stress or any of their combinations. The present sequence
 CC is used in the exemplification of the invention
 XX
 SQ Sequence 146 AA;
 Query Match 97.5%; Score 39; DB 7; Length 146;
 Best Local Similarity 85.7%; Pred. No. 1.3e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGIGDGG 7
 Db 48 GGVGDGG 54
 RESULT 6
 ABB40107
 ID ABB40107 standard; peptide; 157 AA.
 XX
 AC ABB40107;
 XX
 DT 04-FEB-2002 (first entry)
 XX
 DE Peptide #7613 encoded by human foetal liver single exon probe.
 XX
 KW Human; foetal liver; gene expression; single exon nucleic acid probe.

XX Homo sapiens.
 OS
 PN WO200157277-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US000669.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-483447/52.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human fetal liver.
 XX
 PS Claim 27; SEQ ID NO 32742; 639pp + Sequence Listing; English.
 XX
 CC The invention relates to a single exon nucleic acid probe for measuring
 CC human gene expression in a sample derived from human foetal liver. The
 CC single exon nucleic acid probes may be used for predicting, measuring and
 CC displaying gene expression in samples derived from human fetal liver. The
 CC present sequence is a peptide encoded by a single exon nucleic acid probe
 CC of the invention. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 157 AA;
 Query Match 97.5%; Score 39; DB 4; Length 157;
 Best Local Similarity 85.7%; Pred. No. 1.4e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGIGDGG 7
 Db 111 GGVGDGG 117
 RESULT 7
 AAM33743
 ID AAM33743 standard; protein; 157 AA.
 XX
 AC AAM33743;
 XX
 DT 17-OCT-2001 (first entry)
 XX
 DE Peptide #7780 encoded by probe for measuring placental gene expression.
 XX
 KW Probe; microarray; human; placenta; antenatal diagnosis;
 KW genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200157272-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US000663.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488997/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX
XX Claim 27; SEQ ID NO 34012; 654pp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENPs;
CC see AAI31315-AAI57546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders
XX
XX SQ Sequence 157 AA;
XX
XX Query Match 97.5%; Score 39; DB 4; Length 157;
XX Best Local Similarity 85.7%; Pred. No. 1.4e+02;
XX Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 GGIGDGG 7
XX
XX DB 111 GGVGDGG 117
XX
XX
XX RESULT 8
XX AAM73550
XX ID AAM73550 standard; protein; 157 AA.
XX
XX AC AAM73550;
XX
XX DT 06-NOV-2001 (first entry)
XX
XX DE Human bone marrow expressed probe encoded protein SEQ ID NO: 33856.
XX
XX KW Human; bone marrow expressed exon; gene expression analysis; probe;
XX microarray; cancer; leukaemia; lymphoma; myeloma.
XX
XX OS Homo sapiens.
XX
XX PN WO200157276-A2.
XX
XX PD 09-AUG-2001.
XX
XX PF 30-JAN-2001; 2001WO-US000668.
XX
XX PR 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488990/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human bone marrow.
XX
XX Example 4; SEQ ID NO 33856; 658pp + Sequence Listing; English.
XX

CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention
XX
XX SQ Sequence 157 AA;
XX
XX Query Match 97.5%; Score 39; DB 4; Length 157;
XX Best Local Similarity 85.7%; Pred. No. 1.4e+02;
XX Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 GGIGDGG 7
XX
XX DB 111 GGVGDGG 117
XX
XX
XX RESULT 9
XX AAM60865
XX ID AAM60865 standard; protein; 157 AA.
XX
XX AC AAM60865;
XX
XX DT 05-NOV-2001 (first entry)
XX
XX DE Human brain expressed single exon probe encoded protein SEQ ID NO: 32970.
XX
XX KW Human; brain expressed exon; gene expression analysis; probe; microarray;
XX Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.
XX
XX OS Homo sapiens.
XX
XX PN WO200157275-A2.
XX
XX PD 09-AUG-2001.
XX
XX PF 30-JAN-2001; 2001WO-US000667.
XX
XX PR 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains.
XX
XX Example 4; SEQ ID NO 32970; 650pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX epilepsy and cancers. The present sequence is a protein encoded by one of
XX the probes of the invention
XX
XX SQ Sequence 157 AA;
XX
XX Query Match 97.5%; Score 39; DB 4; Length 157;
XX Best Local Similarity 85.7%; Pred. No. 1.4e+02;
XX Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 GGIGDGG 7
XX

Db 111 GGVGDGG 117
||:||||
RESULT 10
ABG55277
ID ABG55277 standard; peptide; 157 AA.
AC ABG55277;
XX 25-FEB-2003 (first entry)
XX Human liver peptide, SEQ ID No 33925.
XX Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
KW hypercholesterolaemia; coronary heart disease.
XX Homo sapiens.
XX WO200157273-A2.
XX 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US000664.
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
(MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488998/53.
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human adult liver.
XX Claim 27; SEQ ID NO 33925; 658bp; English.
XX The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult liver.
CC (i) may be used for predicting, measuring and displaying gene expression
CC in samples derived from human adult liver. The genes identified may be
CC involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC associated with coronary heart disease. ABG47348-ABG59930 represent human
CC liver single exon encoded peptides of the invention. Note: The sequence
CC information for this patent does not appear in the printed specification
CC but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX Sequence 157 AA;
Query Match 97.5%; Score 39; DB 4; Length 157;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGVGDGG 7
||:||||
Db 111 GGVGDGG 117
RESULT 11
ABG43413
ID ABG43413 standard; peptide; 157 AA.
AC ABG43413;
XX 19-AUG-2002 (first entry)
XX Human peptide encoded by genome-derived single exon probe SEQ ID 33078.
XX Human; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; lymphangioleiomyomatosis; Karagener syndrome;
KW pulmonary histiocytosis; fibrocystic pulmonary dysplasia;
KW primary alveolar proteinosis; fibrocystic pulmonary hypertension;
KW hyaline membrane disease.
XX Homo sapiens.
XX WO200186003-A2.
XX 15-NOV-2001.
XX 30-JAN-2001; 2001WO-US000665.
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
(MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2002-114183/15.
XX Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples.
XX Claim 27; SEQ ID NO 33078; 634pp; English.
XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of probes
CC ; the novel set of probes which hybridise at high stringency to a nucleic
CC acid expressed in the human lung; measuring gene expression in a sample
CC derived from human lung, comprising (a) contacting the array with a
CC collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of the
CC array; identifying exons in a eukaryotic genome, comprising (a)
CC algorithmically predicting at least one exon from genomic sequences of
CC the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene expression
CC analysis, and for identifying exons in a gene, particularly using human
CC lung derived mRNA and for the study of lung diseases such as asthma, lung
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,

CC tuberos scleriosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
CC present sequence is a peptide/protein encoded by a single exon probe of
CC the invention. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 157 AA;

Query Match 97.5%; Score 39; DB 5; Length 157;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGIGDGG 7
||:||||
DB 111 GGVGDDG 117

RESULT 12

ABM96534
ID ABM96534 standard; protein; 233 AA.

XX
AC ABM96534;

XX
DT 02-JUN-2005 (first entry)

XX
DE M. xanthus protein sequence, seq id 15733.

XX
KW Transgenic plant; DNA replication; gene regulation; gene expression.

XX
OS Myxococcus xanthus.

XX
PN US6833447-B1.

XX
PD 21-DEC-2004.

XX
PF 10-JUL-2001; 2001US-00902540.

XX
PR 10-JUL-2000; 2000US-0217883P.

XX
PA (MONS) MONSANTO TECHNOLOGY LLC.

XX
PI Goldman BS, Hinkle GJ, Slater SC, Wiegand RC;

XX
DR WPI; 2005-028716/03.

XX
PT New substantially purified Myxococcus xanthus nucleic acid molecule
PT encoding a nitrite reductase, useful for determining gene expression,
PT identifying mutations in a gene of interest, and for constructing
PT mutations in a gene of interest.

XX
PS Example 2; SEQ ID NO 15733; 25pp; English.

XX
CC The invention relates to a substantially purified nucleic acid molecule
CC encoding a nitrite reductase of SEQ ID NO. 11926. Further disclosed is a
CC recombinant DNA construct for expression of a nitrite reductase gene in a
CC plant cell, and a plant cell comprising the recombinant DNA construct.
CC The nucleic acid is useful for determining gene expression, identifying
CC mutations in a gene of interest, and for constructing mutations in a gene
CC of interest. Sequences given in records for SEQ IDs 9692-16825 represent
CC a group of 7134 Myxococcus xanthus proteins and peptides. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from USPTO

XX
SQ Sequence 233 AA;

Query Match 97.5%; Score 39; DB 9; Length 233;
Best Local Similarity 85.7%; Pred. No. 2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGIGDGG 7
||:||||
DB 156 GGVGDDG 162

RESULT 13
ADM20064
ID ADM20064 standard; protein; 239 AA.

XX
AC ADM20064;

XX
DT 20-MAY-2004 (first entry)

XX
DE Protein encoded by novel human channel/transporter gene #132 clone 2.

XX
KW immunosuppressive; antiarthritic; antirheumatic; antiproliferative;
KW cytotstatic; cardiant; vasotropic; cerebroprotective; nootropic;
KW neuroprotective; antibacterial; virucide; fungicide; ophthalmological;
KW gene therapy; channel/transporter protein; rheumatoid arthritis;
KW neoplasm; cardiac arrest; cerebrovascular disorder; cerebral ischemia;
KW angiogenesis; nervous system disorder; Alzheimer's disease;
KW ocular disorder; corneal infection; wound healing;
KW epithelial cell proliferation; skin aging; sunburn; transplantation;
KW chemotaxis; food additive.

XX
OS Homo sapiens.

XX
PN WO200154472-A2.

XX
PD 02-AUG-2001.

XX
PF 17-JAN-2001; 2001WO-US001307.

XX
PR 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 14-AUG-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225577P.

PR 14-AUG-2000; 2000US-0225758P.

PR 14-AUG-2000; 2000US-0225759P.

PR 18-AUG-2000; 2000US-0226279P.

PR 22-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-0226686P.

PR 22-AUG-2000; 2000US-0227182P.

PR 23-AUG-2000; 2000US-0227009P.

PR 30-AUG-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.

PR 01-SEP-2000; 2000US-0229343P.

PR 01-SEP-2000; 2000US-0229344P.

PR 01-SEP-2000; 2000US-0229345P.

PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0233297P.
PR 14-SEP-2000; 2000US-0233398P.
PR 14-SEP-2000; 2000US-0233399P.
PR 14-SEP-2000; 2000US-0233400P.
PR 14-SEP-2000; 2000US-0233401P.
PR 14-SEP-2000; 2000US-0233633P.
PR 14-SEP-2000; 2000US-0233643P.
PR 14-SEP-2000; 2000US-0233653P.
PR 14-SEP-2000; 2000US-0233654P.
PR 14-SEP-2000; 2000US-0233655P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-024617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.

PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-476159/51.
XX N-PSDB; ADM19585.
DR
DR Isolated nucleic acid molecule encoding a channel/transporter protein is
PT used in preventing, treating or ameliorating a medical condition.
XX
XX Claim 11; SEQ ID NO 871; 809pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule encoding a
CC channel/transporter protein or sequences at least 95% identical to a
CC these. The nucleic acids and proteins encoded by them are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC in diagnosing a pathological condition or susceptibility to a
CC pathological condition. The antibodies to the proteins can also be used
CC in alleviating symptoms associated with the disorders and in diagnostic
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays
CC (ELISA). Disorders which are diagnosed or treated include autoimmune
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC arrest, cerebrovascular disorders e.g. cerebral ischemia, angiodenesis,
CC nervous system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection.
CC The polypeptides can also be used to aid wound healing and epithelial
CC cell proliferation, to prevent skin aging due to sunburn, to maintain
CC organs before transplantation, for supporting cell culture of primary
CC tissues, to regenerate tissues and in chemotaxis. The polypeptides can
CC also be used as a food additive or preservative to increase or decrease
CC storage capabilities. This sequence corresponds to a protein of the
XX invention.
XX
SQ Sequence 239 AA;

Query Match 97.5%; Score 39; DB 4; Length 239;
Best Local Similarity 85.7%; Pred. NO. 2.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
||:||||
Db 52 GGVDGG 58

RESULT 14
AAB58809
ID AAB58809 standard; protein; 250 AA.


```

XX AAB58809;
XX
XX 27-MAR-2001 (first entry)
XX
XX Breast and ovarian cancer associated antigen protein sequence SEQ ID 517.
XX
XX Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
XX neotropic; neuroprotective; antiviral; antiallergic; hepatotropic;
XX antidiabetic; antiinflammatory; antiulcer; vulnery; anticonvulsant;
XX antibacterial; antifungal; antiparasitic; cardiant; immune disorder;
XX Addison's disease; allergy; autoimmune haemolytic anaemia;
XX autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
XX multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
XX cardiovascular disorder; wound healing; neurological disease.
XX
XX Homo sapiens.
XX
XX WO200055173-A1.
XX
XX 21-SEP-2000.
XX
XX 08-MAR-2000; 2000WO-US005881.
XX
XX 12-MAR-1999; 99US-0124270P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
XX
XX WPI; 2000-611515/58.
XX
XX N-PSDB; AAF21712.
XX
XX New human breast and ovarian cancer associated gene sequences and the
XX polypeptides encoded by these genes, useful in the prevention, treatment
XX and diagnosis of cancer, immune disorders, cardiovascular disorders and
XX neurological diseases.
XX
XX Claim 11; Page 953; 1299pp; English.
XX
XX Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
XX proteins AAB58711 - AAB59128. The DNA and protein sequences are
XX associated with breast and ovarian cancer. Included in the invention are
XX sequences AAF22032 - AAF22040 and AAB59129 which are used in the
XX isolation and characterisation of the DNA and protein sequences of the
XX invention. The breast and ovarian cancer associated DNA, protein, agonist
XX or antagonist sequences exhibit cytostatic; immunosuppressive; neotropic;
XX neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic;
XX antiinflammatory; antiulcer; vulnery; anticonvulsant; antibacterial;
XX antifungal; antiparasitic and cardiant activity. The polynucleotide and
XX protein sequences are used in the diagnosis of cancer, particularly
XX breast and ovarian cancer. The nucleic acid sequences, proteins, agonists
XX and agonists may also be used in the diagnosis, prevention and treatment
XX of immune disorders e.g. Addison's disease, allergies, autoimmune
XX haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
XX disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
XX cardiovascular disorders such as myocardial ischaemias; wound healing;
XX neurological diseases such as cerebral anoxia and epilepsy; and
XX infectious diseases
XX
XX Sequence 250 AA;
XX
XX Query Match 97.5%; Score 39; DB 3; Length 250;
XX Best Local Similarity 85.7%; Pred. No. 2.2e+02;
XX Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 1 GGIGDGG 7
XX ||:||||
XX 171 GGVGDGG 177
XX
XX Db
XX
XX RESULT 15
XX AAB586774

```

```

ID ABM86774 standard; protein; 326 AA.
XX
XX AC ABM86774;
XX
XX DT 02-JUN-2005 (first entry)
XX
XX DE Rice abiotic stress responsive polypeptide SEQ ID NO:5020.
XX
XX KW abiotic stress tolerance; transgenic plant; plant; cereal; agriculture.
XX
XX OS Oryza sativa.
XX
XX FN WO2003008540-A2.
XX
XX PD 30-JAN-2003.
XX
XX PF 21-JUN-2002; 2002WO-US019668.
XX
XX PR 22-JUN-2001; 2001US-0300112P.
XX
XX PR 24-AUG-2001; 2001US-0314662P.
XX
XX PR 26-SEP-2001; 2001US-0325277P.
XX
XX PR 21-NOV-2001; 2001US-0332132P.
XX
XX PA (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
XX PI Kreps J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F;
XX PI Moughamer T, Provart N, Ricke D, Zhu T;
XX
XX DR WPI; 2003-248011/24.
XX
XX New stress-responsive nucleic acid, useful for altering the
XX responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold
XX stress, salt stress or osmotic stress.
XX
XX Claim 1; SEQ ID NO 5020; 89pp; English.
XX
XX The invention relates to novel abiotic stress responsive polynucleotides
XX and polypeptides. Also disclosed are vectors, expression cassettes, host
XX cells, and plants containing such polynucleotides. Also disclosed are
XX methods for using the polynucleotides and polypeptides to alter the
XX responsiveness of a plant to abiotic stress. The invention is useful in
XX agriculture. The nucleic acid is useful for determining whether a test
XX plant has been exposed to an abiotic stress condition. It is also useful
XX for selecting an agent that alters abiotic stress regulated
XX polynucleotide expression in a plant cell, and to identify a homolog or
XX ortholog to an abiotic stress responsive polynucleotide. The nucleic acid
XX molecule and the polypeptide encoded by it are useful in altering the
XX responsiveness of a plant to an abiotic stress, such as cold stress, salt
XX stress, osmotic stress or any of their combinations. The present sequence
XX is used in the exemplification of the invention
XX
XX SQ Sequence 326 AA;
XX
XX Query Match 97.5%; Score 39; DB 7; Length 326;
XX Best Local Similarity 85.7%; Pred. No. 2.8e+02;
XX Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 1 GGIGDGG 7
XX ||:||||
XX 6 GGVGDGG 12
XX
XX Db
XX
XX RESULT 16
XX ABG28641
XX ID ABG28641 standard; protein; 399 AA.
XX
XX AC ABG28641;
XX
XX DT 18-FEB-2002 (first entry)
XX
XX DE Novel human diagnostic protein #28632.
XX
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

```

KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.

XX WO200175067-A2.

PN 11-OCT-2001.

PD 30-MAR-2001; 2001WO-US008631.

PF 31-MAR-2000; 2000US-00540217.

XX 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

PA Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS92828.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

XX Claim 20; SEQ ID NO 59000; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 399 AA;

Query Match 97.5%; Score 39; DB 4; Length 399;
Best Local Similarity 85.7%; Pred. No. 3.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7

DB 28 GGVGDGG 34

RESULT 17

ADM19814

ID ADM19814 standard; protein; 445 AA.

XX ADM19814;

XX 20-MAY-2004 (first entry)

DT Protein encoded by novel human channel/transporter gene #132.

XX immunosuppressive; antiarthritic; antirheumatic; antiproliferative;
XX cytotactic; cardiant; vasotropic; cerebroprotective; nootropic;
KW neuroprotective; antibacterial; virucide; fungicide; ophthalmological;
KW

KW gene therapy; channel/transporter protein; rheumatoid arthritis;
KW neoplasm; cardiac arrest; cerebrovascular disorder; cerebral ischemia;
KW angiogenesis; nervous system disorder; Alzheimer's disease;
KW ocular disorder; corneal infection; wound healing;
KW epithelial cell proliferation; skin aging; sunburn; transplantation;
KW chemotaxis; food additive.
XX
OS Homo sapiens.

XX WO200154472-A2.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US001307.

XX 31-JAN-2000; 2000US-0179065P.

XX 04-FEB-2000; 2000US-0180628P.

XX 24-FEB-2000; 2000US-0184664P.

XX 02-MAR-2000; 2000US-0186350P.

XX 16-MAR-2000; 2000US-0189874P.

XX 17-MAR-2000; 2000US-0190078P.

XX 18-APR-2000; 2000US-0198123P.

XX 19-MAY-2000; 2000US-0205515P.

XX 07-JUN-2000; 2000US-0209467P.

XX 28-JUN-2000; 2000US-0214886P.

XX 30-JUN-2000; 2000US-0215135P.

XX 07-JUL-2000; 2000US-0216647P.

XX 07-JUL-2000; 2000US-0216880P.

XX 11-JUL-2000; 2000US-0217487P.

XX 11-JUL-2000; 2000US-0217496P.

XX 14-JUL-2000; 2000US-0218290P.

XX 26-JUL-2000; 2000US-0220963P.

XX 26-JUL-2000; 2000US-0220964P.

XX 14-AUG-2000; 2000US-0224518P.

XX 14-AUG-2000; 2000US-0224519P.

XX 14-AUG-2000; 2000US-0225213P.

XX 14-AUG-2000; 2000US-0225214P.

XX 14-AUG-2000; 2000US-0225266P.

XX 14-AUG-2000; 2000US-0225267P.

XX 14-AUG-2000; 2000US-0225268P.

XX 14-AUG-2000; 2000US-0225270P.

XX 14-AUG-2000; 2000US-0225447P.

XX 14-AUG-2000; 2000US-0225757P.

XX 14-AUG-2000; 2000US-0225758P.

XX 14-AUG-2000; 2000US-0225759P.

XX 18-AUG-2000; 2000US-0226279P.

XX 22-AUG-2000; 2000US-0226681P.

XX 22-AUG-2000; 2000US-0226686P.

XX 22-AUG-2000; 2000US-0227182P.

XX 23-AUG-2000; 2000US-0227009P.

XX 30-AUG-2000; 2000US-0228924P.

XX 01-SEP-2000; 2000US-0229287P.

XX 01-SEP-2000; 2000US-0229343P.

XX 01-SEP-2000; 2000US-0229344P.

XX 01-SEP-2000; 2000US-0229345P.

XX 05-SEP-2000; 2000US-0229509P.

XX 05-SEP-2000; 2000US-0229513P.

XX 06-SEP-2000; 2000US-0230437P.

XX 08-SEP-2000; 2000US-0231242P.

XX 08-SEP-2000; 2000US-0231243P.

XX 08-SEP-2000; 2000US-0231244P.

XX 08-SEP-2000; 2000US-0231413P.

XX 08-SEP-2000; 2000US-0231414P.

XX 08-SEP-2000; 2000US-0232080P.

XX 12-SEP-2000; 2000US-0231968P.

XX 14-SEP-2000; 2000US-0232397P.

XX 14-SEP-2000; 2000US-0232398P.

XX 14-SEP-2000; 2000US-0232399P.

XX 14-SEP-2000; 2000US-0232400P.

XX 14-SEP-2000; 2000US-0232401P.

XX 14-SEP-2000; 2000US-0233063P.

PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234423P.
PR 21-SEP-2000; 2000US-0234424P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0244674P.
PR 08-NOV-2000; 2000US-0244759P.
PR 08-NOV-2000; 2000US-0244766P.
PR 08-NOV-2000; 2000US-0244777P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.

PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-476159/51.
DR N-PSDB; ADM19335.
XX
XX
PT Isolated nucleic acid molecule encoding a channel/transporter protein is
PT used in preventing, treating or ameliorating a medical condition.
XX
XX Claim 11; SEQ ID NO 621; 809pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule encoding a
CC channel/transporter protein or sequences at least 95% identical to a
CC these. The nucleic acids and proteins encoded by them are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC in diagnosing a pathological condition or susceptibility to a
CC pathological condition. The antibodies to the proteins can also be used
CC in alleviating symptoms associated with the disorders and in diagnostic
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays
CC (ELISA). Disorders which are diagnosed or treated include autoimmune
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC arrest, cerebrovascular disorders e.g. cerebral ischemia, angiodenesis,
CC nervous system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection.
CC The polypeptides can also be used to aid wound healing and epithelial
CC cell proliferation, to prevent skin aging due to sunburn, to maintain
CC organs before transplantation, for supporting cell culture of primary
CC tissues, to regenerate tissues and in chemotaxis. The polypeptides can
CC also be used as a food additive or preservative to increase or decrease
CC storage capabilities. This sequence corresponds to a protein of the
CC invention.
XX
SQ Sequence 445 AA;
Query Match 97.5%; Score 39; DB 4; Length 445;
Best Local Similarity 85.7%; Pred. No. 3.7e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGIGDGG 7
Db 52 GGVGDDG 58
RESULT 18
ABP43927
ID ABP43927 standard; protein; 525 AA.
XX ABP43927;
XX
XX 26-FEB-2003 (first entry)
DT
XX
XX FLJ20261 fis clone COLF7630.
DE
XX
XX Neuroprotective; immunomodulator; cancer; chromosome 17; cytostatic;
KW anti-inflammatory; gene therapy; nutritional supplement; wound; burn;
KW ulcer; Alzheimer's disease; Huntington's disease;
KW amyotrophic lateral sclerosis; autoimmune disorder; inflammation;
KW vulnery.
XX
XX Homo sapiens.
OS
XX
XX WO200231111-A2.
PN
XX
XX 18-APR-2002.

XX PF 11-OCT-2001; 2001WO-US027760.
XX PR 12-OCT-2000; 2000US-00687527.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
XX PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX DR N-PSDB; AB061171.
XX DR WPI: 2002-426278/45.
XX PR N-PSDB; AB061171.
XX PT New polypeptides and their encoded proteins, useful as nutritional
XX PT sources or supplements, or in gene therapy, particularly for treating
XX PT wounds, Alzheimer's disease, amyotrophic lateral sclerosis, cancer or
XX PT inflammation.
XX PS Claim 20; SEQ ID # 830; 357pp + Sequence Listing; English.
XX CC The invention relates to 446 newly isolated polynucleotide sequences. The
XX CC activity of polynucleotides of the invention may be described as,
XX CC vulnary, neuroprotective, immunomodulator, cytostatic and anti-
XX CC inflammatory. Compositions comprising nucleic acids of the invention are
XX CC useful for treating a mammalian subject, or as nutritional sources or
XX CC supplements. These are useful in gene therapy, particularly for treating
XX CC wounds, burns or ulcers, Alzheimer's disease, Huntington's disease,
XX CC amyotrophic lateral sclerosis, autoimmune disorders, cancer or
XX CC inflammation. The nucleic acids and polypeptides are also useful in
XX CC diagnostic and research methods. The sequences given in records ABP43544-
XX CC ABP43989 represent polypeptides encoded by polynucleotides of the
XX CC invention. NOTE: The sequence data for this patent did not form part of
XX CC the printed specification, but was obtained in electronic format directly
XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 525 AA;

Query Match 97.5%; Score 39; DB 5; Length 525;
Best Local Similarity 85.7%; Pred. No. 4.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
Db 128 GGVGDGG 134
||:||||

RESULT 19
ABO55390
ID ABO55390 standard; protein; 37 AA.
XX AC ABO55390;
XX AC ABO55390;
XX DT 29-JUL-2004 (first entry)
XX DE Human genome derived single exon protein #1624.
XX KW Human; gene expression; single exon probe; microarray;
XX KW alternative splicing event; genomic alteration.
XX OS Homo sapiens.
XX PN US2003194704-A1.
XX PD 16-OCT-2003.
XX PF 03-APR-2002; 2002US-00029386.
XX PR 03-APR-2002; 2002US-00029386.
XX PA (PENN/) PENN S G.
XX PA (RANK/) RANK D R.
XX PA (HANZ/) HANZEL D K.
XX XX

PI Penn SG, Rank DR, Hanzel DK;
XX WPI: 2004-119264/12.
XX PT New human genome-derived single exon nucleic acid probes useful for human
XX PT gene expression analysis, for identifying or characterizing alternative
XX PT splicing events, for assessing genomic alterations or as tools for
XX PT surveying tissues.
XX PS Claim 45; SEQ ID NO 29024; 80pp; English.
XX CC The invention relates to a nucleic acid probe for measuring human gene
XX CC expression, comprising any of the 27,400 fully defined nucleotide
XX CC sequences in the specification, or their complements or fragments, and
XX CC encoding at least 8 amino acids of any of the 6888 amino acid sequences
XX CC fully defined in the specification. The probe is a single exon probe that
XX CC hybridises under high stringency conditions to a nucleic acid molecule
XX CC expressed in human cells or tissues. Also included are a spatially-
XX CC addressable set of single exon nucleic acid probes for measuring human
XX CC gene expression (comprising a plurality of single exon nucleic acid
XX CC probes cited above, where each of the plurality of probes is, separately
XX CC and addressably isolatable or amplifiable from the plurality), a single
XX CC exon microarray for measuring human gene expression, a method of
XX CC measuring human gene expression, a vector comprising the single exon
XX CC probe cited above, an ORF-encoded peptide comprising at least 8
XX CC contiguous amino acids of any of the above-mentioned amino acid
XX CC sequences (optionally with conservative amino acid substitutions), an
XX CC isolated antibody that binds specifically to a peptide cited above,
XX CC methods of selling and/or licensing single exon probes or microarrays to
XX CC a customer desiring to measure gene expression, a method of providing
XX CC human gene expression data by subscription, and a computer-readable
XX CC storage medium which contains a database having a plurality of records
XX CC (each record including data on the expression of a single exon probe
XX CC cited above. The probe, methods and apparatus are useful in gene
XX CC expression analysis. The probes may be used as tools for surveying
XX CC tissues to detect the presence of expressed messages that contain their
XX CC specific exon, or in constructing genome-derived single exon microarrays.
XX CC In addition, the probes are used in identifying and characterising
XX CC alternative splicing events, in detecting and characterising gross
XX CC alterations in the genomic locus that includes their exon, in assessing
XX CC smaller genomic alterations, in priming the synthesis of nucleic acids,
XX CC or in expressing the ORF-encoded peptide. The present sequence is a human
XX CC single exon probe protein of the invention. Note: The sequence data for
XX CC this patent did not form part of the printed specification, but was
XX CC obtained in electronic format directly from USPTO at
XX CC seqdata.uspto.gov/sequence.html?DocID=20030194704
XX SQ Sequence 37 AA;

Query Match 95.0%; Score 38; DB 8; Length 37;
Best Local Similarity 85.7%; Pred. No. 50;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
Db 31 GGVGDGG 37
||:||||

RESULT 20
ABB61028
ID ABB61028 standard; protein; 427 AA.
XX AC ABB61028;
XX AC ABB61028;
XX DT 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster polypeptide SEQ ID NO 9876.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical.
XX OS Drosophila melanogaster.
XX XX

PN WO200171042-A2.
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US009231.
 XX
 PR 23-MAR-2000; 2000US-0191637P.
 PR 11-JUL-2000; 2000US-00614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI; 2001-656860/75.
 DR N-PSDB; ABL05131.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signaling and cell-cell
 PT interactions.
 XX
 PS Disclosure; SEQ ID NO 9876; 21pp + Sequence Listing; English.
 CC
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signaling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS57737-
 CC AB872072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 427 AA;
 Query Match 95.0%; Score 38; DB 4; Length 427;
 Best Local Similarity 85.7%; Pred. NO. 5.2e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GGIGDGG 7
 Db 296 GGLGDGG 302
 ||:||||
 ||:||||
 RESULT 21
 ABO83915
 ID ABO83915 standard; protein; 430 AA.
 XX
 AC ABO83915;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE Pseudomonas aeruginosa polypeptide #16090.
 XX
 KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
 XX
 OS Pseudomonas aeruginosa.
 XX
 XX US6551795-B1.
 PN
 PD 22-APR-2003.
 XX
 PF 18-FEB-1999; 99US-00252991.
 XX
 PR 18-FEB-1998; 98US-0074788P.
 PR 27-JUL-1998; 98US-0094190P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
 XX
 DR WPI; 2003-615309/58.
 DR N-PSDB; ABD17486.

XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
 PT useful as molecular targets for diagnostics, prophylaxis and treatment of
 PT pathological conditions resulting from bacterial infection.
 XX
 PS Disclosure; SEQ ID NO 32661; 455pp; English.
 XX
 CC The invention relates to Pseudomonas aeruginosa polypeptides and the
 CC polynucleotides encoding them. The sequences are useful in diagnosis and
 CC therapy of pathological conditions, as molecular targets for diagnostics,
 CC prophylaxis and treatment of pathological conditions resulting from a
 CC bacterial infection, for evaluating a compound, such as a polypeptide,
 CC for the ability to bind a P. aeruginosa nucleic acid, as components of
 CC effective antibacterial targets, as targets for antibacterial drugs,
 CC including anti-P. aeruginosa drugs, as templates for recombinant
 CC production of P. aeruginosa-derived peptides or polypeptides, as target
 CC components for diagnosis and/or treatment of P. aeruginosa-caused
 CC infection, and in detection of P. aeruginosa sequences or other sequences
 CC of Pseudomonas species using biochip technology. Sequences ABO67826-
 CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format from USPTO at
 CC seqdata.uspto.gov/sequence.html
 XX
 SQ Sequence 430 AA;
 Query Match 95.0%; Score 38; DB 7; Length 430;
 Best Local Similarity 85.7%; Pred. NO. 5.2e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GGIGDGG 7
 Db 64 GGLGDGG 70
 ||:||||
 ||:||||
 RESULT 22
 AAU02902
 ID AAU02902 standard; protein; 682 AA.
 XX
 AC AAU02902;
 XX
 DT 12-SEP-2001 (first entry)
 XX
 DE Angiotensin converting enzyme (ACEV) splice variant protein #2.
 XX
 KW Angiotensin converting enzyme splice variant; ACEV; interleukin 6;
 KW granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
 KW platelet-derived endothelial cell growth factor; cardiovascular disease;
 KW cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C;
 KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
 KW myocardial infarction; coronary arterial thrombosis; renal disease;
 KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;
 KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;
 KW nonaroidotic pulmonary granulomatous disease; endothelial abnormality;
 KW vascular disorder; asbestosis.
 XX
 OS Homo sapiens.
 XX
 PN WO200136632-A2.
 XX
 PD 25-MAY-2001.
 XX
 PF 17-NOV-2000; 2000WO-IL000766.
 XX
 PR 17-NOV-1999; 99IL-00132978.
 PR 10-DEC-1999; 99IL-00133455.
 XX
 PA (COMP-) COMPUGEN LTD.
 XX
 PI Levine Z, David A, Azar I, Khosravi R, Bernstein J;
 XX
 DR WPI; 2001-336004/35.
 DR N-PSDB; AAS06002.

XX Novel alternative splicing variants e.g. variant of angiotensin
FT converting enzyme (ACEV), useful in identifying candidate compounds
FT capable of binding to the variant and to detect anti-variant antibodies.
XX
XX Claim 4; Fig 2; 519pp; English.
XX
XX The sequence represents an angiotensin converting enzyme splice variant
CC (ACEV) polypeptide. The polypeptides of the invention include variants of
CC granulocyte colony stimulating factor receptor, glucagon, interleukin 6,
CC platelet-derived endothelial cell growth factor, cyclin-dependent kinase
CC inhibitor 1C, cellular tumour antigen p53, and vasoactive intestinal
CC polypeptide receptor 2. The polypeptides and their associated nucleic
CC acids are useful for identification of variant sequences and detection of
CC candidate compounds capable of binding to the molecules. The sequences of
CC the invention can be used in the treatment and diagnosis of various
CC disorders including cardiovascular diseases such as arteriosclerosis,
CC myocardial infarction and coronary arterial thrombosis, renal diseases
CC such as diabetic nephropathy, muscular diseases such as hypertrophy,
CC immune disorders such as immune complex nephritis, multiple sclerosis,
CC cancer, sarcoidosis, nonrheumatoid pulmonary granulomatous diseases such
CC as asbestosis and vascular pathologies involving an endothelial
CC abnormality such as deep vein thrombosis
XX
XX Sequence 682 AA;
SQ

Query Match 90.0%; Score 36; DB 4; Length 682;
Best Local Similarity 85.7%; Pred. NO. 1.7e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGIGDGG 7
DB 508 GGIGEGG 514
|||:|

RESULT 23
AAE10129
ID AAE10129 standard; protein; 11096 AA.
XX
XX AAE10129;
XX
XX 29-NOV-2001 (first entry)
XX
XX Streptomyces noursei nystatin gene, NysC.
DE
XX Polyketide synthase; PKS; macrolide; nystatin; PKS gene cluster;
KW antifungal; antibiotic; PKS type I.
XX
XX Streptomyces noursei.
XX
XX Key Location/Qualifiers
FH Domain 35..455
FT /label= KS3 domain
FT /note= "Ketoreductase (KS) domain"
FT Domain 546..858
FT /label= AT3 domain
FT /note= "Acyltransferase (AT) domain"
FT Domain 872..1073
FT /label= DH3 domain
FT /note= "Dehydratase (DH) domain"
FT Domain 1381..1628
FT /label= KR3 domain
FT /note= "Ketoreductase (KR) domain"
FT Domain 1662..1735
FT /label= ACP3 domain
FT /note= "Acyl carrier protein (ACP) domain"
FT Domain 1757..2180
FT /label= KS4 domain
FT /note= "Ketoreductase (KS) domain"
FT Domain 2291..2603
FT /label= AT4 domain
FT /note= "Acyltransferase (AT) domain"
FT Domain 2617..2818

/label= DH4 domain
/note= "Dehydratase (DH) domain"
3124..3371
/label= KR4 domain
/note= "Ketoreductase (KR) domain"
3407..3480
/label= ACP4 domain
/note= "Acyl carrier protein (ACP) domain"
3501..3924
/label= KS5 domain
/note= "Ketoreductase (KS) domain"
4032..4346
/label= AT5 domain
/note= "Acyltransferase (AT) domain"
4360..4561
/label= DH5 domain
/note= "Dehydratase (DH) domain"
4953..5239
/label= ER5 domain
/note= "Enoylreductase (ER) domain"
5248..5495
/label= KR5 domain
/note= "Ketoreductase (KR) domain"
5528..5601
/label= ACP5 domain
/note= "Acyl carrier protein (ACP) domain"
5623..5646
/label= KS6 domain
/note= "Ketoreductase (KS) domain"
6166..6478
/label= AT6 domain
/note= "Acyltransferase (AT) domain"
6492..6704
/label= DH6 domain
/note= "Dehydratase (DH) domain"
7038..7281
/label= KR6 domain
/note= "Ketoreductase (KR) domain"
7315..7388
/label= ACP6 domain
/note= "Acyl carrier protein (ACP) domain"
7408..7831
/label= KS7 domain
/note= "Ketoreductase (KS) domain"
7939..8253
/label= AT7 domain
/note= "Acyltransferase (AT) domain"
8267..8470
/label= DH7 domain
/note= "Dehydratase (DH) domain"
8812..9086
/label= KR7 domain
/note= "Ketoreductase (KR) domain"
9120..9193
/label= ACP7 domain
/note= "Acyl carrier protein (ACP) domain"
9214..9637
/label= KS8 domain
/note= "Ketoreductase (KS) domain"
9758..10072
/label= AT8 domain
/note= "Acyltransferase (AT) domain"
10086..10289
/label= DH8 domain
/note= "Dehydratase (DH) domain"
10657..10904
/label= KR8 domain
/note= "Ketoreductase (KR) domain"
10939..11012
/label= ACP8 domain
/note= "Acyl carrier protein (ACP) domain"
XX
XX WO200159126-A2.
PN

```

XX PD 16-AUG-2001.
XX PF 08-FEB-2001; 2001WO-CB000509.
XX PR 08-FEB-2000; 2000GB-00002840.
XX PR 10-APR-2000; 2000GB-00008786.
XX PR 14-APR-2000; 2000GB-00009387.
XX PA (UYNO-) UNIV NORGES TEKNISK NATURVITENSKAPELIGE.
XX PA (SNTF) SINTEF STIFTELSEN IND TEK FORSK.
XX PA (ALPH-) ALPHARMA AS.
XX PA (SINV-) SINVENT AS.
XX PA (DZIE/) DZIEGLEWSKA H.
XX PA (ZOTC/) ZOTCHEV S B.
XX PA (SEKU/) SEKUROVA O N.
XX PA (FJAE/) FJAERVIK E.
XX PA (BRAU/) BRAUTASET T.
XX PA (STRO/) STROM A R.
XX PA (VALL/) VALLA S.
XX PI Zotchev SB, Sekurova ON, Fjaervik E, Brautaset T, Strom AR;
XX PI Valla S, Ellingsen TE, Sletta H, Gulliksen O;
XX PR WPI: 2001-557614/62.
XX DR N-PSDB; AADI7184.
XX CC New nystatin polyketide synthase polynucleotides and polypeptides, useful
XX PT as antibiotics and antifungals.
XX PS Claim 15; Page 170-176; 266pp; English.
XX CC The present invention relates to the cloning and sequencing of the gene
XX CC cluster encoding a modular type I polyketide synthase (PKS) enzyme
XX CC involved in the biosynthesis of the macrolide antibiotic nystatin. The
XX CC nystatin PKS is useful as antifungal antibiotics. The present sequence is
XX CC a PKS type I encoding Streptomyces noursei nystatin gene, Nysc
XX SQ Sequence 11096 AA;
XX Query Match 90.0%; Score 36; DB 4; Length 11096;
XX Best Local Similarity 85.7%; Pred. No. 2.4e+04;
XX Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX QY 1 GGIGDGG 7
XX Db 8425 GGFDDGG 8431
XX RESULT 24
XX AAY54970
XX ID AAY54970 standard; peptide; 19 AA.
XX AC AAY54970;
XX XX
XX DT 15-FEB-2000 (first entry)
XX DE Peptide ligand for fibrin polymerisation site.
XX KW Thrombus imaging; fibrin polymerisation site; technetium-99m; Tc-99m;
XX KW GPIIb/IIIa receptor; cyclic peptide ligand.
XX OS Synthetic.
XX PH Key Location/Qualifiers
XX FT Modified-site 1
XX FT /note= "cysteine residue blocked by covalent linkage to
XX FT an acetamido group"
XX FT Modified-site 3
XX FT /note= "cysteine residue blocked by covalent linkage to
XX FT an acetamido group"
XX PN US5968476-A.

```

```

XX PD 19-OCT-1999.
XX PF 07-JUN-1995; 95US-00484773.
XX PR 21-MAY-1992; 92US-00886052.
XX PR 11-JUL-1994; 94US-00273274.
XX PA (DIAT-) DIATIDE INC.
XX PI Dean RT, Lister-James J;
XX DR WPI: 2000-021733/02.
XX PT A complex used for thrombus imaging comprises technetium-99m complexed
XX PT with a peptide ligand for GPIIb/IIIa receptor.
XX PS Example 2; Col 13-14; 18pp; English.
XX CC This sequence represents a peptide ligand for the fibrin polymerisation
XX CC site. The invention relates to a complex (A) for thrombus imaging
XX CC comprises technetium-99m (Tc-99m) completed with a reagent comprising a
XX CC peptide (P) with 4 to 100 amino acids sequence and a Tc-99m binding
XX CC moiety covalently bound to (P). (P) is selected from a linear peptide
XX CC ligand for a GPIIb/IIIa receptor not comprising the amino acid sequence
XX CC (arginine-glycine-aspartate), a peptide ligand for a polymerisation site
XX CC of fibrin, and a cyclic peptide ligand for the GPIIb/IIIa receptor. The
XX CC thrombus imaging reagents provided by the present invention can be used
XX CC for visualising thrombi in a mammalian body when Tc-99m is labelled
XX SQ Sequence 19 AA;
XX Query Match 87.5%; Score 35; DB 3; Length 19;
XX Best Local Similarity 85.7%; Pred. No. 79;
XX Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX QY 1 GGIGDGG 7
XX Db 9 GGTGDGG 15
XX RESULT 25
XX AAM16732
XX ID AAM16732 standard; protein; 61 AA.
XX AC AAM16732;
XX XX
XX DT 12-OCT-2001 (first entry)
XX DE Peptide #3166 encoded by probe for measuring cervical gene expression.
XX KW Probe; human; microarray; gene expression; cervical epithelial cell;
XX KW cervical cancer.
XX OS Homo sapiens.
XX PN WO200157278-A2.
XX XX
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000670.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX XX (MOLE-) MOLECULAR DYNAMICS INC.
XX PA Penn SG, Hanzel DK, Chen W, Rank DR;
XX PI

```

XX WPI; 2001-488901/53.
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human cervical epithelial cells.
XX

PS Claim 27; SEQ ID NO 21558; 487bp; English.

XX The present invention relates to human single exon nucleic acid probes
CC (SENPs: see AAI10068-AAI2459). The present sequence is a peptide encoded
CC by one such probe. The SENPs are derived from human Hela cells. The SENPs
CC can be used to produce a single exon microarray, which can be used for
CC measuring human gene expression in a sample derived from human cervical
CC epithelial cells. By measuring gene expression, the probes are therefore
CC useful in grading and/or staging of diseases of the cervix, notably
CC cervical cancer. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 61 AA;

Query Match 87.5%; Score 35; DB 4; Length 61;
Best Local Similarity 85.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
|||
Db 32 GGYGDGG 38

RESULT 26
ABB35718
ID ABB35718 standard; peptide; 61 AA.

XX ABB35718;

XX 04-FEB-2002 (first entry)

DE Peptide #3224 encoded by human foetal liver single exon probe.

XX Human; foetal liver; gene expression; single exon nucleic acid probe.

OS Homo sapiens.

XX WO200157277-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000669.

XX 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483447/52.

XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human fetal liver.
XX Claim 27; SEQ ID NO 28353; 639pp + Sequence Listing; English.

XX The invention relates to a single exon nucleic acid probe for measuring
CC human gene expression in a sample derived from human foetal liver. The
CC single exon nucleic acid probes may be used for predicting, measuring and
CC displaying gene expression in samples derived from human fetal liver. The

CC present sequence is a peptide encoded by a single exon nucleic acid probe
CC of the invention. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 61 AA;

Query Match 87.5%; Score 35; DB 4; Length 61;
Best Local Similarity 85.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
|||
Db 32 GGYGDGG 38

RESULT 27
ABB34279
ID ABB34279 standard; peptide; 61 AA.

XX ABB34279;

XX 04-FEB-2002 (first entry)

DE Peptide #1785 encoded by human foetal liver single exon probe.

XX Human; foetal liver; gene expression; single exon nucleic acid probe.

XX Homo sapiens.

XX WO200157277-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000669.

XX 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483447/52.

XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human fetal liver.

XX Claim 27; SEQ ID NO 26914; 639pp + Sequence Listing; English.

XX The invention relates to a single exon nucleic acid probe for measuring
CC human gene expression in a sample derived from human foetal liver. The
CC single exon nucleic acid probes may be used for predicting, measuring and
CC displaying gene expression in samples derived from human fetal liver. The
CC present sequence is a peptide encoded by a single exon nucleic acid probe
CC of the invention. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 61 AA;

Query Match 87.5%; Score 35; DB 4; Length 61;
Best Local Similarity 85.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
|||
Db 32 GGYGDGG 38


```
RESULT 28
AAM29220
ID AAM29220 standard; protein; 61 AA.
XX
XX AAM29220;
AC AAM29220;
DT 17-OCT-2001 (first entry)
XX
XX Peptide #3257 encoded by probe for measuring placental gene expression.
DE
DE Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder.
KW
XX Homo sapiens.
OS
XX WO200157272-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US000663.
PF
XX
XX 04-FEB-2000; 2000US-0180312P.
PR
XX 26-MAY-2000; 2000US-0207456P.
PR
XX 30-JUN-2000; 2000US-00608408.
PR
XX 03-AUG-2000; 2000US-00632366.
PR
XX 21-SEP-2000; 2000US-0234687P.
PR
XX 27-SEP-2000; 2000US-0236359P.
PR
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488897/53.
XX
XX Claim 27; SEQ ID NO 29489; 654pp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENPs;
XX see AAI31315-AAI57546). The present sequence is a peptide encoded by one
XX such probe. The probes are useful for producing a microarray for
XX predicting, measuring and displaying gene expression in samples derived
XX from human placenta. The probes are useful for antenatal diagnosis of
XX human genetic disorders
XX
XX Query Match 87.5%; Score 35; DB 4; Length 61;
XX Best Local Similarity 85.7%; Pred. No. 2.4e+02;
XX Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 GGIGDGG 7
XX |||||
XX Db 32 GGYGDGG 38
XX
XX RESULT 29
XX ABB29120
XX ID ABB29120 standard; peptide; 61 AA.
XX
XX ABB29120;
AC ABB29120;
DT 01-FEB-2002 (first entry)
XX
XX Peptide #1771 encoded by breast cell single exon nucleic acid probe.
DE
DE Human; microarray; single exon probe; gene expression; breast; disease;
KW cancer.
KW
XX Homo sapiens.
XX
```

```
OS Homo sapiens.
XX
XX WO200157271-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US000662.
PF
XX
XX 04-FEB-2000; 2000US-0180312P.
PR
XX 26-MAY-2000; 2000US-0207456P.
PR
XX 30-JUN-2000; 2000US-00608408.
PR
XX 03-AUG-2000; 2000US-00632366.
PR
XX 21-SEP-2000; 2000US-0234687P.
PR
XX 27-SEP-2000; 2000US-0236359P.
PR
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-496933/54.
XX
XX New spatially-addressable set of single exon nucleic acid probes, useful
XX for measuring gene expression in sample derived from human breast,
XX comprises number of single exon nucleic acid probes.
XX
XX Claim 27; SEQ ID NO 12088; 327pp + Sequence Listing; English.
XX
XX The invention relates to a spatially-addressable set of single exon
XX nucleic acid probes for measuring gene expression in a sample derived
XX from human breast and BT 474 cells. The method involves contacting the
XX probes with a collection of detectably labelled nucleic acids derived
XX from mRNA of human breast, and then measuring the label bound to each
XX probe of the microarray. The probes are useful for verifying the
XX expression of regions of genomic DNA predicted to encode proteins. They
XX are useful for gene discovery, and for determining predisposition and/or
XX prognosing breast disease. Gene expression analysis is useful for
XX assessing the toxicity of chemical agents on cells. The microarray of
XX this invention presents a far greater diversity of probes for measuring
XX gene expression, with far less bias than expressed sequence tag
XX microarrays. The method is suitable for rapid production of functional
XX information from genomic sequence. The present sequence is a peptide
XX encoded by a single exon nucleic acid probe of the invention. Note: The
XX sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences
XX
XX Query Match 87.5%; Score 35; DB 4; Length 61;
XX Best Local Similarity 85.7%; Pred. No. 2.4e+02;
XX Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 GGIGDGG 7
XX |||||
XX Db 32 GGYGDGG 38
XX
XX RESULT 30
XX ABB30553
XX ID ABB30553 standard; peptide; 61 AA.
XX
XX ABB30553;
AC ABB30553;
XX
XX 01-FEB-2002 (first entry)
DT
XX
XX Peptide #3204 encoded by breast cell single exon nucleic acid probe.
DE
DE Human; microarray; single exon probe; gene expression; breast; disease;
KW cancer.
KW
XX Homo sapiens.
XX
```


PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488900/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human bone marrow.
XX
XX Example 4; SEQ ID NO 27773; 658pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention
XX
XX Sequence 61 AA;
SQ

Query Match 87.5%; Score 35; DB 4; Length 61;
Best Local Similarity 85.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GGIGDGG 7
Db 32 GGYGDGG 38

RESULT 33
AAM68909
ID AAM68909 standard; protein; 61 AA.
XX
XX AAM68909;
XX
XX 06-NOV-2001 (first entry)
XX Human bone marrow expressed probe encoded protein SEQ ID NO: 29215.
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 29215.
XX
XX Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma.
XX
XX Homo sapiens.
OS
XX WO200157276-A2.
PN
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US000668.
PF
XX
XX 04-FEB-2000; 2000US-0180312P.
PR
XX 26-MAY-2000; 2000US-0207456P.
PR
XX 30-JUN-2000; 2000US-00608408.
PR
XX 03-AUG-2000; 2000US-00632366.
PR
XX 21-SEP-2000; 2000US-0234687P.
PR
XX 27-SEP-2000; 2000US-0236359P.
PR
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI; 2001-488900/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human bone marrow.
XX
XX Example 4; SEQ ID NO 29215; 658pp + Sequence Listing; English.
XX

CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention
XX
XX Sequence 61 AA;
SQ

Query Match 87.5%; Score 35; DB 4; Length 61;
Best Local Similarity 85.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GGIGDGG 7
Db 32 GGYGDGG 38

RESULT 34
AAM55074
ID AAM55074 standard; protein; 61 AA.
XX
XX AAM55074;
XX
XX 05-NOV-2001 (first entry)
XX Human brain expressed single exon probe encoded protein SEQ ID NO: 27179.
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 27179.
XX
XX Human; brain expressed exon; gene expression analysis; probe; microarray;
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.
XX
XX Homo sapiens.
OS
XX WO200157275-A2.
PN
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US000667.
PF
XX
XX 04-FEB-2000; 2000US-0180312P.
PR
XX 26-MAY-2000; 2000US-0207456P.
PR
XX 30-JUN-2000; 2000US-00608408.
PR
XX 03-AUG-2000; 2000US-00632366.
PR
XX 21-SEP-2000; 2000US-0234687P.
PR
XX 27-SEP-2000; 2000US-0236359P.
PR
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
PT brains.
XX
XX Example 4; SEQ ID NO 27179; 650pp + Sequence Listing; English.
PS
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention
XX
XX Sequence 61 AA;
SQ

Query Match 87.5%; Score 35; DB 4; Length 61;
Best Local Similarity 85.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GGIGDGG 7

PD 09-AUG-2001.
XX 29-JAN-2001; 2001WO-US000661.
PF XX
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-476286/51.
PI Novel single exon nucleic acid probe used to measuring gene expression in
DR a human breast.
XX Claim 27; SEQ ID NO 11776; 322pp; English.
XX The present invention relates to novel single exon nucleic acid probes
CC (see AAI00010-AA110067). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for measuring human gene expression in
CC a human breast sample, where the probe hybridises at high stringency to a
CC nucleic acid expressed in the human breast. The probes are useful for
CC predicting, diagnosing, grading, staging, monitoring and prognosing
CC diseases of the human breast, particularly those diseases with polygenic
CC aetiology. The diseases include: breast cancer; disorders of development,
CC inflammatory diseases of the breast, fibrocystic changes, proliferative
CC breast disease and non-carcinoma tumours. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 61 AA;
Query Match 87.5%; Score 35; DB 4; Length 61;
Best Local Similarity 85.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GGIGDGG 7
Db |||||
32 GGYGDGG 38
RESULT 38
AA0449
ID AAM04449 standard; protein; 61 AA.
XX
XX AAM04449;
AC
XX 09-OCT-2001 (first entry)
DT
DE Peptide #3131 encoded by probe for measuring breast gene expression.
XX Probe; human; breast disease; breast cancer; development disorder;
KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.
KW
OS Homo sapiens.
XX
XX WO200157270-A2.
PN
XX 09-AUG-2001.
PD
XX 29-JAN-2001; 2001WO-US000661.
PF
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR

PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI; 2001-476286/51.
DR Novel single exon nucleic acid probe used to measuring gene expression in
PT a human breast.
PT
XX Claim 27; SEQ ID NO 13189; 322pp; English.
XX The present invention relates to novel single exon nucleic acid probes
CC (see AAI00010-AA110067). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for measuring human gene expression in
CC a human breast sample, where the probe hybridises at high stringency to a
CC nucleic acid expressed in the human breast. The probes are useful for
CC predicting, diagnosing, grading, staging, monitoring and prognosing
CC diseases of the human breast, particularly those diseases with polygenic
CC aetiology. The diseases include: breast cancer; disorders of development,
CC inflammatory diseases of the breast, fibrocystic changes, proliferative
CC breast disease and non-carcinoma tumours. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 61 AA;
Query Match 87.5%; Score 35; DB 4; Length 61;
Best Local Similarity 85.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GGIGDGG 7
Db |||||
32 GGYGDGG 38
RESULT 39
ABG38490
ID ABG38490 standard; peptide; 61 AA.
XX
XX ABG38490;
AC
XX 19-AUG-2002 (first entry)
DT
DE Human peptide encoded by genome-derived single exon probe SEQ ID 28155.
XX
XX Human; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangiomyomatosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease.
XX
OS Homo sapiens.
XX
XX WO200186003-A2.
PN
XX 15-NOV-2001.
PD
XX 30-JAN-2001; 2001WO-US000665.
PF
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR

PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2002-114183/15.
XX Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples.
XX Claim 27; SEQ ID NO 28155; 634pp; English.
XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of probes
CC ; the novel set of probes which hybridise at high stringency to a nucleic
CC acid expressed in the human lung; measuring gene expression in a sample
CC derived from human lung, comprising (a) contacting the array with a
CC collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of the
CC array; identifying exons in a eukaryotic genome, comprising (a)
CC algorithmically predicting at least one exon from genomic sequences of
CC the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene expression
CC analysis, and for identifying exons in a gene, particularly using human
CC lung derived mRNA and for the study of lung diseases such as asthma, lung
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
CC present sequence is a peptide/protein encoded by a single exon probe of
CC the invention. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX Sequence 61 AA;
SQ
Query Match 87.5%; Score 35; DB 5; Length 61;
Best Local Similarity 85.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGGDGG 7
|||
DB 32 GGYDGG 38
RESULT 40
ID ABP33054
XX ABP33054 standard; protein; 137 AA.
XX AC ABP33054;
XX 09-JUL-2002 (first entry)
XX Human ORF2027 protein, SEQ ID NO:4054.
DE

XX Human; ORF; open reading frame; ORFX; drug screening; diagnosis;
KW disease monitoring; cytokine; cell proliferation; cell differentiation;
KW immune modulation; haematopoiesis regulation; tissue growth;
KW angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; tumour inhibition; bodily characteristic; fertility;
KW behaviour; cancer; proliferative disorder; neurological disorder;
KW cardiovascular disease; immune system disorder; organ transplantation;
KW tissue growth disorder; tissue regeneration disease; diabetes mellitus;
KW hypothyroidism; cholesterol ester storage disease; infection; vulvular;
KW vasotropic; antipsoriatic; antidiabetic; cytostatic; nootropic;
KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
KW cardiant; hypotensive; antithyroid; antiinflammatory; immunomodulator;
KW dermatological; analgesic; virucide; antibacterial; fungicide.
XX Homo sapiens.
OS
XX WO200190366-A2.
PN
XX 29-NOV-2001.
PD
XX 24-MAY-2001; 2001WO-US017076.
PF
XX 24-MAY-2000; 2000US-0206690P.
PR
XX (CURA-) CURAGEN CORP.
PA
XX Leach MD, Shimkets RA;
PI
XX WPI; 2002-106200/14.
DR
XX N-PSDB; ABN77080.
DR
XX Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and disorders related to organ
PT transplantation.
XX Claim 10; Page 1271; 2508pp; English.
PS
XX Sequences ABP31028-ABP3561 represent 4534 novel human proteins
CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-
CC ABN79587 represent cDNAs encoding them. The invention also encompasses
CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
CC referred to as ORFX) proteins, polynucleotides at least 85% identical to
CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
CC polynucleotides, the recombinant production of ORFX proteins, antibodies
CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and
CC polypeptides, methods of screening for modulators of ORFX expression or
CC activity, and methods of screening individuals for a predisposition to an
CC ORFX-associated disorder. The ORFX proteins of the invention have a wide
CC range of biological activities, such as cytokine, cell proliferation,
CC cell differentiation, immune modulation, haematopoiesis regulation,
CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
CC chemokinetic activity, haemostatic activity, thrombolytic activity,
CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,
CC and antiinfective activity, and may also be involved in the determination
CC of bodily characteristics, fertility and behaviour. ORFX proteins,
CC nucleic acids and antibodies may be used in the treatment of cancers,
CC other proliferative disorders such as psoriasis and benign tumours,
CC neurological disorders such as epilepsy and Alzheimer's disease,
CC cardiovascular diseases, immune system disorders, disorders related to
CC organ transplantation, disorders of tissue growth and regeneration, ester
CC storage disease, and infectious diseases caused by viral, bacterial,
CC fungal and other pathogens. ORFX nucleic acids may also be used as a
CC source of primers and probes, in the detection of ORFX genomic sequences
CC or transcripts, in the identification and cloning of homologous
CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
CC nucleic acids may additionally be used to produce transgenic animals
CC which may be useful for studying the function and/or activity of ORFX
CC protein, and in drug screening. The ORFX proteins may also be used as
CC immunogens to generate specific antibodies, which are useful in the
CC diagnosis, treatment and monitoring of ORFX-associated diseases

XX SQ Sequence 137 AA;
Query Match 87.5%; Score 35; DB 5; Length 137;
Best Local Similarity 85.7%; Pred. No. 5.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGIGDGG 7
|| |||||
Db 56 GGTGDGG 62
RESULT 41
ID ABO61865
XX ABO61865 standard; protein; 142 AA.
XX AC ABO61865;
XX DT 29-JUL-2004 (first entry)
XX DE Klebsiella pneumoniae polypeptide segid 8382.
XX KW Recombinant expression vector; transcription regulatory element;
XX KW Klebsiella pneumoniae protein; antibacterial; vaccine.
XX OS Klebsiella pneumoniae.
XX PN US6610836-B1.
XX PD 26-AUG-2003.
XX PF 27-JAN-2000; 2000US-00489039.
XX PR 29-JAN-1999; 99US-0117747P.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PI Breton GL, Osborne M;
XX WPI; 2003-895346/82.
XX DR N-PSDB; ACH95416.
XX PT New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
XX PT preparing a vaccine composition against Klebsiella pneumoniae.
XX PS Disclosure; SEQ ID NO 8382; 932pp; English.
XX CC The invention describes a new isolated nucleic acid encoding a Klebsiella
XX CC pneumoniae polypeptide. Also described are: a recombinant expression
XX CC vector comprising the nucleic acid, operably linked to a transcription
XX CC regulatory element; and a cell comprising the recombinant expression
XX CC vector. The nucleic acid is useful for preparing a vaccine composition
XX CC against Klebsiella pneumoniae. This is the amino acid sequence of a
XX CC Klebsiella pneumoniae polypeptide of the invention
XX SQ Sequence 142 AA;
Query Match 87.5%; Score 35; DB 7; Length 142;
Best Local Similarity 85.7%; Pred. No. 5.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGIGDGG 7
|| |||||
Db 95 GGAGDGG 101
RESULT 42
ID ADI42042
XX ADI42042 standard; protein; 165 AA.
XX AC ADI42042;
XX XX
XX DT 22-APR-2004 (first entry)

XX DE Plant transcription factor #237.
XX KW transgenic; plant; enhanced tolerance to abiotic stress;
KW phosphate tolerance; hormone sensitivity; disease resistance;
KW sugar sensing; flowering; flower structure; stem bifurcation;
KW branching pattern; apical dominance; trichome; stem morphology;
KW root growth; root hair; seed development; cell proliferation;
KW cell differentiation; premature senescence; necrosis; plant size;
KW leaf morphology; seed morphology; seed biochemistry; root anthocyanin;
KW plant anthocyanin; light response; shade avoidance; bioinformatic;
KW transcription factor; gene; ds.
XX Oryza sativa.
XX US2004019927-A1.
XX PN 29-JAN-2004.
XX PD 25-FEB-2003; 2003US-00374780.
XX PF 18-APR-2001; 2001US-00837944.
XX PR (SHER/) SHERMAN B K.
XX PA (RIEC/) RIECHMANN J L.
XX PA (JIAN/) JIANG C.
XX PA (HEAR/) HEARD J B.
XX PA (HAKE/) HAAKE V.
XX PA (CREE/) CREELMAN R A.
XX PA (RATC/) RATCLIFFE O.
XX PA (ADAM/) ADAM L J.
XX PA (REUB/) REUBER T L.
XX PA (KEDD/) KEDDIE J.
XX PA (BROU/) BROUN P E.
XX PA (PILG/) PILGRIM M L.
XX PA (DUBE/) DUBELL A N.
XX PA (PINE/) PINEDA O.
XX PA (YUGG/) YU G.
XX Sherman BK, Riechmann JL, Jiang C, Heard JE, Haake V;
PI Creelman RA, Ratcliffe O, Adam LJ, Reuber TU, Keddie J, Broun PE;
PI Pilgrim ML, Dubell AN, Pineda O, Yu G;
XX WPI; 2004-132245/13.
XX PT New transgenic plant comprising a recombinant polynucleotide of any one
XX PT of more than 500 nucleotide sequences, useful in bioinformatic search
XX PT methods.
XX Claim 1; SEQ ID NO 505; 435pp; English.
XX CC The invention describes a transgenic plant comprising a recombinant
XX CC polynucleotide of any one of more than 500 nucleotide sequences fully
XX CC defined in the specification or its complement. The method of the
XX CC invention can be used to produce a plant having altered traits such as:
XX CC enhanced tolerance to abiotic stress; phosphate tolerance; hormone
XX CC sensitivity; disease resistance; sugar sensing; early or late flowering;
XX CC altered flower structure; change in stem bifurcations, altered branching
XX CC pattern, reduced apical dominance, reduced trichome density; lack of
XX CC trichomes; reduced ectopic trichome development; altered trichome
XX CC development; increase in trichome number; altered stem morphology;
XX CC increased root growth; increased root hairs; altered seed development;
XX CC altered cell proliferation or cell differentiation; rapid development;
XX CC premature senescence; increased necrosis; increase in seedling or plant
XX CC size; decreased plant size; leaf morphology; seed morphology; seed
XX CC biochemistry; increase in root anthocyanins; increase in plant
XX CC anthocyanins, or alteration in light response or shade avoidance. The
XX CC transgenic plant, polynucleotides and polypeptides are useful in
XX CC bioinformatic search methods. This is the amino acid sequence of a plant
XX CC transcription factor, and an orthologue of Arabidopsis thaliana
XX CC creation of a transgenic plant with altered traits.

SQ Sequence 165 AA;

Query Match 87.5%; Score 35; DB 8; Length 165;
 Best Local Similarity 85.7%; Pred. No. 6.2e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
 |||||
 Db 65 GGAGDGG 71

RESULT 43
 ADO62236
 ID ADO62236 standard; protein; 165 AA.
 XX
 AC ADO62236;
 XX
 DT 15-JUL-2004 (first entry)
 XX
 DE Transcription factor G47 orthologous sequence, SEQ ID 703.
 XX
 KW Plant; transcription factor; transgenic plant; abiotic stress tolerance;
 KW osmotic stress tolerance; cold tolerance; heat tolerance;
 KW low nitrogen tolerance; low phosphate tolerance; fungal disease;
 KW glyphosate resistance; flowering; fertility; seed development.
 XX
 OS Oryza sativa.
 XX
 PN WO2004031349-A2.
 XX
 PD 15-APR-2004.
 XX
 PF 18-SEP-2003; 2003WO-US030292.
 XX
 PR 18-SEP-2002; 2002US-0411837P.
 PR 17-DEC-2002; 2002US-0434166P.
 PR 24-APR-2003; 2003US-0465809P.
 XX
 PA (MEND-) MENDEL BIOTECHNOLOGY INC.
 XX
 PI Jiang C, Heard JE, Ratcliffe O, Creelman RA, Adam LJ, Reuber TL;
 PI Riechmann JL, Haake V, Dubell AN, Keddle JS, Sherman BK;
 XX
 DR WPI; 2004-330163/30.
 XX
 PT New recombinant polynucleotide encoding transcription factor
 PT polypeptides, useful for producing transgenic plants with advantageous
 PT properties compared to a reference plant.
 XX
 PS Claim 1; SEQ ID NO 703; 510pp; English.
 XX
 CC The present invention relates to novel plant transcription factor
 CC proteins (I) and nucleotide sequences (II) (ADO61534-ADO63778). The
 CC sequences can be used to produce transgenic plants, which overexpress
 CC (II), where the transgenic plant has an altered trait as compared to a
 CC non-transgenic plant or wild-type plant. The transgenic plant comprises
 CC an altered trait selected from increased tolerance to abiotic stress,
 CC increased tolerance to osmotic stress, increased tolerance to cold,
 CC increased germination in cold, increased tolerance to heat, increased
 CC germination in heat, increased tolerance to freezing conditions,
 CC increased tolerance to low nitrogen conditions, increased tolerance to
 CC low phosphate conditions, increased tolerance to disease, including
 CC fungal disease and particularly Erysiphe, Fusarium and Botrytis,
 CC increased tolerance to multiple fungal pathogens, increased resistance to
 CC glyphosate, increased sensitivity to ABA, reduced sensitivity to ABA,
 CC increased sensitivity to ACC, altered sugar sensing, increased tolerance
 CC to sugars, altered carbon/nitrogen sensing, early flowering, late
 CC flowering, altered flower structure, loss of flower determinacy, reduced
 CC fertility, altered shoot meristem development, altered branching pattern,
 CC altered stem morphology, altered vascular tissue structure, reduced
 CC apical dominance, altered trichome density, altered trichome development,
 CC altered trichome structure, altered root development, altered shade
 CC avoidance, altered seed development, altered seed ripening, altered seed

CC germination, slow growth, fast growth, altered cell differentiation,
 CC altered cell proliferation, altered cell expansion, altered phase change,
 CC altered senescence, abnormal embryo development, altered programmed cell
 CC death, lethality when overexpressed, altered necrosis patterns, increased
 CC plant size, increased biomass, large seedlings, dwarfed plants, dark
 CC green leaves, change in leaf shape, increased leaf size and mass, light
 CC green or gray leaves, glossy leaves, altered abaxial/adaxial polarity,
 CC altered seed coloration, altered seed size, altered seed shape, large
 CC seed, increased leaf wax, increased leaf fatty acids, altered seed oil
 CC content, altered seed protein content, altered seedprenyl content,
 CC altered leaf prenyl lipid content, increased anthocyanin levels, and
 CC decreased anthocyanin levels. Note: the sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 165 AA;

Query Match 87.5%; Score 35; DB 8; Length 165;
 Best Local Similarity 85.7%; Pred. No. 6.2e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
 |||||
 Db 65 GGAGDGG 71

RESULT 44
 ADO63760
 ID ADO63760 standard; protein; 228 AA.
 XX
 AC ADO63760;
 XX
 DT 15-JUL-2004 (first entry)
 XX
 DE Transcription factor G3644 orthologous sequence, SEQ ID 2227.
 XX
 KW Plant; transcription factor; transgenic plant; abiotic stress tolerance;
 KW osmotic stress tolerance; cold tolerance; heat tolerance;
 KW low nitrogen tolerance; low phosphate tolerance; fungal disease;
 KW glyphosate resistance; flowering; fertility; seed development.
 XX
 OS Oryza sativa.
 XX
 PN WO2004031349-A2.
 XX
 PD 15-APR-2004.
 XX
 PF 18-SEP-2003; 2003WO-US030292.
 XX
 PR 18-SEP-2002; 2002US-0411837P.
 PR 17-DEC-2002; 2002US-0434166P.
 PR 24-APR-2003; 2003US-0465809P.
 XX
 PA (MEND-) MENDEL BIOTECHNOLOGY INC.
 XX
 PI Jiang C, Heard JE, Ratcliffe O, Creelman RA, Adam LJ, Reuber TL;
 PI Riechmann JL, Haake V, Dubell AN, Keddle JS, Sherman BK;
 XX
 DR WPI; 2004-330163/30.
 XX
 PT New recombinant polynucleotide encoding transcription factor
 PT polypeptides, useful for producing transgenic plants with advantageous
 PT properties compared to a reference plant.
 XX
 PS Disclosure; SEQ ID NO 2227; 510pp; English.
 XX
 CC The present invention relates to novel plant transcription factor
 CC proteins (I) and nucleotide sequences (II) (ADO61534-ADO63778). The
 CC sequences can be used to produce transgenic plants, which overexpress
 CC (II), where the transgenic plant has an altered trait as compared to a
 CC non-transgenic plant or wild-type plant. The transgenic plant comprises

an altered trait selected from increased tolerance to abiotic stress, increased tolerance to osmotic stress, increased tolerance to cold, increased germination in cold, increased tolerance to heat, increased germination in heat, increased tolerance to freezing conditions, increased tolerance to low nitrogen conditions, increased tolerance to low phosphate conditions, increased tolerance to disease, including fungal disease and particularly Erysiphe, Fusarium and Botrytis, increased tolerance to multiple fungal pathogens, increased resistance to glyphosate, increased sensitivity to ABA, reduced sensitivity to ABA, increased sensitivity to ACC, altered sugar sensing, increased tolerance to sugars, altered carbon/nitrogen sensing, early flowering, late flowering, altered flower structure, loss of flower determinacy, reduced fertility, altered shoot meristem development, altered branching pattern, altered stem morphology, altered vascular tissue structure, reduced apical dominance, altered trichome density, altered trichome development, altered trichome structure, altered root development, altered shade avoidance, altered seed development, altered seed ripening, altered seed germination, slow growth, fast growth, altered cell differentiation, altered cell proliferation, altered cell expansion, altered phase change, altered senescence, abnormal embryo development, altered programmed cell death, lethality when overexpressed, altered necrosis patterns, increased plant size, increased biomass, large seedlings, dwarfed plants, dark green leaves, change in leaf shape, increased leaf size and mass, light green or gray leaves, glossy leaves, altered abaxial/adaxial polarity, altered seed coloration, altered seed size, altered seed shape, large seed, increased leaf wax, increased leaf fatty acids, altered seed oil content, altered seed protein content, altered seedprenyl content, altered leaf prenyl lipid content, increased anthocyanin levels, and decreased anthocyanin levels. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 228 AA;

Query Match 87.5%; Score 35; DB 8; Length 228;
Best Local Similarity 85.7%; Pred. No. 8.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGIGDGG 7
DB 104 GGAGDGG 110

RESULT 45
AEA26249
ID AEA26249 standard; protein; 228 AA.

XX AEA26249;

AC AEA26249;

DT 28-JUL-2005 (first entry)

DE Stress tolerant plant-related transcription factor protein SeqID90.

XX transcription factor; transgenic plant; agriculture; drought resistance;
KW stress tolerance.

XX Oryza sativa.

OS WO2005047516-A2.

XX 26-MAY-2005.

PD 12-NOV-2004; 2004WO-US037584.

PF 13-NOV-2003; 2003US-00714887.

PR 05-DEC-2003; 2003US-0527658P.

PR 05-FEB-2004; 2004US-0542928P.

XX (MEND-) MENDEL BIOTECHNOLOGY INC.

XX Heard JE, Riechmann JL, Creelman RA, Ratcliffe OJ, Canales RD;
PI Repetti P, Kumimoto RW, Gutterson NI, Reuber TL, Pineda O;

PI Sherman BK, Morrison TA, Keddle JS, Jiang C, Century KS, Adam L;
PI Zhang JZ, Hempel FD, Libby JW;
XX WPI; 2005-372386/38.
DR N-PSDB; AEA26248.
XX New transgenic plants for producing commercially or agriculturally useful plants having improved tolerance to drought, shade and low nitrogen conditions.

PS Example 8; SEQ ID NO 90; 407pp; English.

XX This invention relates to a novel plant transcription factor polypeptides, the DNA sequences which encode them and their use in creating transgenic plants. The transgenic plant and methods are useful for producing commercially or agriculturally useful plants having improved tolerance to drought, shade and low nitrogen conditions when compared to wild-type reference plants. The present sequence is that of a plant transcription factor protein which was used during the development of the transgenic plants of the invention.

XX Sequence 228 AA;

Query Match 87.5%; Score 35; DB 9; Length 228;
Best Local Similarity 85.7%; Pred. No. 8.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGIGDGG 7
DB 104 GGAGDGG 110

RESULT 46

ADO63770
ID ADO63770 standard; protein; 233 AA.

XX ADO63770;

DT 15-JUL-2004 (first entry)

DE Transcription factor G3651 orthologous sequence, SEQ ID 2237.

XX Plant; transcription factor; transgenic plant; abiotic stress tolerance;
KW osmotic stress tolerance; cold tolerance; heat tolerance;
KW low nitrogen tolerance; low phosphate tolerance; fungal disease;
XX glyphosate resistance; flowering; fertility; seed development.

OS Oryza sativa.

XX WO2004031349-A2.

XX 15-APR-2004.

PF 18-SEP-2003; 2003WO-US030292.

PR 18-SEP-2002; 2002US-0411837P.

PR 17-DEC-2002; 2002US-0434166P.

PR 24-APR-2003; 2003US-0465809P.

XX (MEND-) MENDEL BIOTECHNOLOGY INC.

XX Jiang C, Heard JE, Ratcliffe O, Creelman RA, Adam LJ, Reuber TL;
PI Riechmann JL, Haake V, Dubell AN, Keddle JS, Sherman BK;

XX WPI; 2004-330163/30.

DR N-PSDB; ADO63769.

XX New recombinant polynucleotide encoding transcription factor polypeptides, useful for producing transgenic plants with advantageous properties compared to a reference plant.

XX Disclosure; SEQ ID NO 2237; 510pp; English.

CC The present invention relates to novel plant transcription factor
 CC proteins (I) and nucleotide sequences (II) (AD061534-AD063778). The
 CC sequences can be used to produce transgenic plants, which overexpress
 CC (II), where the transgenic plant has an altered trait as compared to a
 CC non-transgenic plant or wild-type plant. The transgenic plant comprises
 CC an altered trait selected from increased tolerance to abiotic stress,
 CC increased tolerance to osmotic stress, increased tolerance to cold,
 CC increased germination in cold, increased tolerance to heat, increased
 CC germination in heat, increased tolerance to freezing conditions,
 CC increased tolerance to low nitrogen conditions, increased tolerance to
 CC low phosphate conditions, increased tolerance to disease, including
 CC fungal disease and particularly Erysiphe, Fusarium and Botrytis.
 CC increased tolerance to multiple fungal pathogens, increased resistance to
 CC glyphosate, increased sensitivity to ABA, reduced sensitivity to ABA,
 CC increased sensitivity to ACC, altered sugar sensing, increased tolerance
 CC to sugars, altered carbon/nitrogen sensing, early flowering, late
 CC flowering, altered flower structure, loss of flower determinacy, reduced
 CC fertility, altered shoot meristem development, altered branching pattern,
 CC altered stem morphology, altered vascular tissue structure, reduced
 CC apical dominance, altered trichome density, altered trichome development,
 CC altered trichome structure, altered root development, altered shade
 CC avoidance, altered seed development, altered seed ripening, altered seed
 CC germination, slow growth, fast growth, altered cell differentiation,
 CC altered cell proliferation, altered cell expansion, altered phase change,
 CC altered senescence, abnormal embryo development, altered programmed cell
 CC death, lethality when overexpressed, altered necrosis patterns, increased
 CC plant size, increased biomass, large seedlings, dwarfed plants, dark
 CC green leaves, change in leaf shape, increased leaf size and mass, light
 CC green or gray leaves, glossy leaves, altered abaxial/adaxial polarity,
 CC altered seed coloration, altered seed size, altered seed shape, large
 CC seed, increased leaf wax, increased leaf fatty acids, altered seed oil
 CC content, altered seed protein content, altered seedprenyl content,
 CC altered leaf prenyl lipid content, increased anthocyanin levels, and
 CC decreased anthocyanin levels. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 233 AA;

Query Match 87.5%; Score 35; DB 8; Length 233;
 Best Local Similarity 85.7%; Pred. No. 8.6e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
 || |||||
 Db 112 GGAGDGG 118

RESULT 47
 AEA26259
 ID AEA26259 standard; protein; 233 AA.

XX AEA26259;

XX 28-JUL-2005 (first entry)

XX Stress tolerant plant-related transcription factor protein SeqID100.

XX transcription factor; transgenic plant; agriculture; drought resistance;
 XX stress tolerance.

XX Oryza sativa.

XX WO2005047516-A2.

XX 26-MAY-2005.

XX 12-NOV-2004; 2004WO-US037584.

XX 13-NOV-2003; 2003US-00714887.

XX 03-DEC-2003; 2003US-0527658P.

XX 05-FEB-2004; 2004US-0542928P.

XX (MEND-) MENDEL BIOTECHNOLOGY INC.
 PA Heard JE, Riechmann JL, Creelman RA, Ratcliffe OJ, Canales RD;
 XX Repetti P, Kumimoto RW, Guttererson NI, Reuber TL, Pineda O;
 PI Sherman BK, Morrison TA, Keddle JS, Jiang C, Century KS, Adam L;
 PI Zhang JZ, Hempel FD, Libby JM;
 XX WPI; 2005-372386/38.
 DR N-PSDB; AEA26258.

XX New transgenic plants for producing commercially or agriculturally useful
 PT plants having improved tolerance to drought, shade and low nitrogen
 PT conditions.

XX Example 8; SEQ ID NO 100; 407pp; English.

XX This invention relates to a novel plant transcription factor
 CC polypeptides, the DNA sequences which encode them and their use in
 CC creating transgenic plants. The transgenic plant and methods are useful
 CC for producing commercially or agriculturally useful plants having
 CC improved tolerance to drought, shade and low nitrogen conditions when
 CC compared to wild-type reference plants. The present sequence is that of a
 CC plant transcription factor protein which was used during the development
 CC of the transgenic plants of the invention.

XX SQ Sequence 233 AA;

Query Match 87.5%; Score 35; DB 9; Length 233;
 Best Local Similarity 85.7%; Pred. No. 8.6e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
 || |||||
 Db 112 GGAGDGG 118

RESULT 48

ADT60392
 ID ADT60392 standard; protein; 248 AA.

XX ADT60392;

XX 13-JAN-2005 (first entry)

XX Plant polypeptide, SEQ ID 10469.

XX Plant; transgenic; cold tolerance; growth rate; drought tolerance;
 KW disease resistance; galactomannan production; plant growth regulator;
 KW heat tolerance; herbicide tolerance; lignin production;
 KW extreme osmotic condition tolerance; pathogens resistance;
 KW pest resistance; yield improvement; seed oil yield; seed protein yield.

XX Viridiplantae.

XX US2004216190-A1.

XX 28-OCT-2004.

XX 18-DEC-2003; 2003US-00739930.

XX 28-APR-2003; 2003US-00424599.

XX 28-APR-2003; 2003US-00425115.

XX (KOVA/) KOVALIC D K.

XX Kovalic DK;

XX WPI; 2004-757369/74.

XX New recombinant DNA constructs useful in the field of biochemistry and
 PT genetics, and in particular for producing transgenic plants with improved
 PT biological characteristics.

XX PS Claim 2; SEQ ID NO 10469; 14pp; English.

XX CC The invention relates a recombinant DNA construct comprising a

CC polynucleotide having any of 5544 nucleotide sequences (cDNAs SEQ ID NO:

CC 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences

CC (SEQ ID NO: 5545-11098). The cDNAs and proteins are from corn, soybean,

CC Arabidopsis, wheat and rape but the specification does not indicate which

CC sequences is derived from which organism. Also included is a method of

CC producing a plant having an improved property, comprising transforming a

CC plant with a recombinant DNA construct comprising a promoter region

CC functional in a plant cell operably joined to a polynucleotide encoding a

CC polypeptide associated with the property, and growing the transformant

CC plant. The property is selected from improving plant cold tolerance, for

CC manipulating growth rate in plant cells by modification of the cell cycle

CC pathway, for improving plant drought tolerance, for providing increased

CC resistance to plant disease, for galactanmanan production, for production

CC of plant growth regulators, for improving plant heat tolerance, for

CC improving plant tolerance to herbicides, for increasing the rate of

CC homologous recombination in plants, for lignin production, for improving

CC plant tolerance to extreme osmotic conditions, for improving plant

CC tolerance to pathogens or pests, for yield improvement by modification of

CC photosynthesis, for modifying seed oil yield and/or content, for

CC modifying seed protein yield and/or content, for yield improvement by

CC modification of carbohydrate, nitrogen or phosphorus use and/or uptake

CC and for yield improvement by providing improved plant growth and

CC development under at least one stress condition. The polynucleotide may

CC also encode a plant transcription factor. The methods and compositions of

CC the present invention are useful in the field of biochemistry and

CC genetics, in particular for producing transgenic plants with improved

CC biological characteristics such as increased yield, improved nitrogen

CC flow, increasing plant tolerance to cold or heat, improving plant

CC tolerance to extreme osmotic and drought conditions, and improving plant

CC tolerance to plant pests or pathogens. They can also be used in physical

CC arrays of molecules, plant breeding markers, computer-based storage and

CC analysis systems. The present sequence is one of the 5544 plant protein

CC sequences of the invention. Note: The sequence data for this patent did

CC not form part of the printed specification, but was obtained in

CC electronic format directly from USPTO at

CC seqdata.uspto.gov/sequence.html?DocID=20040216190.

XX SQ Sequence 248 AA;

Query Match 87.5%; Score 35; DB 8; Length 248;

Best Local Similarity 85.7%; Pred. NO. 9.2e+02;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGIGDGG 7

Db 18 GGAGDGG 24

RESULT 49

ABG13340

ID ABG13340 standard; protein; 281 AA.

XX AC ABG13340;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #13331.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX FN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX

PR 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

PA Drmanac RT, Liu C, Tang YT;

PI WPI; 2001-639362/73.

XX N-PSDB; AAS77527.

DR New isolated polynucleotide and encoded polypeptides, useful in

XX PT diagnostic, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity.

PS Claim 20; SEQ ID NO 43699; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)

CC sequences. (I) is useful as hybridisation probes, polymerase chain

CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,

CC and in recombinant production of (II). The polynucleotides are also used

CC in diagnostics as expressed sequence tags for identifying expressed

CC genes. (II) is useful in gene therapy techniques to restore normal

CC activity of (II) or to treat disease states involving (II). (II) is

CC useful for generating antibodies against it, detecting or quantitating a

CC polypeptide in tissue, as molecular weight markers and as a food

CC supplement. (II) and its binding partners are useful in medical imaging

CC of sites expressing (II). (I) and (II) are useful for treating disorders

CC involving aberrant protein expression or biological activity. The

CC polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic

CC amino acid sequences of the invention. Note: The sequence data for this

CC patent did not appear in the printed specification, but was obtained in

CC electronic format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 281 AA;

Query Match 87.5%; Score 35; DB 4; Length 281;

Best Local Similarity 85.7%; Pred. NO. 1e+03;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGIGDGG 7

Db 48 GGAGDGG 54

RESULT 50

ABO80780

ID ABO80780 standard; protein; 281 AA.

XX AC ABO80780;

XX DT 29-JUL-2004 (first entry)

XX DE Pseudomonas aeruginosa polypeptide #12955.

XX KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

XX OS Pseudomonas aeruginosa.

XX FN US6551795-B1.

XX PD 22-APR-2003.

XX PF 18-FEB-1999; 99US-00252991.

XX PR 18-FEB-1998; 98US-0074788P.

XX PR 27-JUL-1998; 98US-0094190P.

XX

PA (GENO-) GENOME THERAPEUTICS CORP.
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX WPI; 2003-615309/58.
DR N-PSDB; ABD14351.
XX
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide, of
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
PS Disclosure; SEQ ID NO 29526; 455pp; English.
XX
CC The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
SQ Sequence 281 AA;

Query Match 87.5%; Score 35; DB 7; Length 281;
Best Local Similarity 85.7%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
DB 26 GGAGDGG 32

Search completed: February 27, 2006, 09:45:55
Job time : 191 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: February 27, 2006, 09:50:10 ; Search time 47 Seconds
(without alignments)
12.313 Million cell updates/sec

Title: US-10-027-015B-2

Perfect score: 40

Sequence: 1 GGIGGG 7

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	100.0	449	2	US-09-252-991A-20941, A
2	39	97.5	200	2	US-09-248-796A-24658
3	39	97.5	233	2	US-09-902-540-15733
4	38	95.0	430	2	US-09-252-991A-32661
5	35	87.5	142	2	US-09-489-039A-8382
6	35	87.5	281	2	US-09-252-991A-29526
7	35	87.5	327	2	US-09-270-767-45414
8	35	87.5	334	2	US-09-252-991A-27369
9	35	87.5	341	2	US-09-902-540-16071
10	35	87.5	405	2	US-09-489-039A-7341
11	35	87.5	406	2	US-09-252-991A-24561
12	35	87.5	416	2	US-09-252-991A-22415
13	35	87.5	416	2	US-09-252-991A-23585
14	35	87.5	535	2	US-09-252-991A-31062
15	34	85.0	10	1	US-08-484-905-4
16	34	85.0	10	2	US-08-481-985B-4
17	34	85.0	10	2	US-08-370-476-4
18	34	85.0	13	1	US-08-484-905-6
19	34	85.0	13	2	US-08-481-985B-6
20	34	85.0	13	2	US-08-370-476-6
21	34	85.0	15	1	US-08-484-905-8
22	34	85.0	15	2	US-08-481-985B-8
23	34	85.0	15	2	US-08-370-476-8
24	34	85.0	19	1	US-08-484-905-12
25	34	85.0	19	2	US-08-481-985B-12
26	34	85.0	19	2	US-08-370-476-12
27	34	85.0	21	1	US-08-484-905-14

28	34	85.0	21	2	US-08-481-985B-14	Sequence 14, Appl
29	34	85.0	21	2	US-08-370-476-14	Sequence 14, Appl
30	34	85.0	37	1	US-08-461-597-3	Sequence 3, Appl
31	34	85.0	37	1	US-08-535-298-3	Sequence 3, Appl
32	34	85.0	37	4	PCT-US94-05569A-3	Sequence 3, Appl
33	34	85.0	37	4	PCT-US94-05569-3	Sequence 3, Appl
34	34	85.0	170	2	US-09-248-796A-21808	Sequence 12, Appl
35	34	85.0	256	2	US-09-270-767-42925	Sequence 12, Appl
36	34	85.0	310	2	US-09-498-520A-12	Sequence 10, Appl
37	34	85.0	332	2	US-09-498-520A-10	Sequence 10, Appl
38	34	85.0	332	2	US-09-134-000C-4988	Sequence 4988, Ap
39	34	85.0	341	2	US-09-252-991A-27327	Sequence 27327, A
40	34	85.0	356	2	US-09-252-991A-25656	Sequence 25656, A
41	34	85.0	413	2	US-09-149-476-751	Sequence 751, App
42	34	85.0	434	2	US-09-252-991A-29256	Sequence 29256, A
43	34	85.0	469	2	US-09-252-991A-26584	Sequence 26584, A
44	34	85.0	614	2	US-09-489-039A-9961	Sequence 9961, Ap
45	34	85.0	693	2	US-09-949-016-7243	Sequence 7243, Ap
46	34	85.0	806	2	US-08-999-774A-6	Sequence 6, Appl
47	34	85.0	912	2	US-09-949-016-6786	Sequence 6786, Ap
48	34	85.0	918	2	US-09-252-991A-24163	Sequence 24163, A
49	34	85.0	918	2	US-09-809-920-2	Sequence 2, Appl
50	33	82.5	11	1	US-08-482-880-10	Sequence 10, Appl
51	33	82.5	11	1	US-08-273-274-10	Sequence 10, Appl
52	33	82.5	11	1	US-08-475-041-10	Sequence 10, Appl
53	33	82.5	11	1	US-08-484-773-10	Sequence 10, Appl
54	33	82.5	11	1	US-08-361-864-14	Sequence 14, Appl
55	33	82.5	12	1	US-08-753-781-22	Sequence 22, Appl
56	33	82.5	14	1	US-08-335-832-10	Sequence 10, Appl
57	33	82.5	14	2	US-09-141-127-4	Sequence 4, Appl
58	33	82.5	14	2	US-08-825-765-12	Sequence 12, Appl
59	33	82.5	15	1	US-08-753-781-4	Sequence 4, Appl
60	33	82.5	16	1	US-08-482-880-11	Sequence 11, Appl
61	33	82.5	16	1	US-08-273-274-11	Sequence 11, Appl
62	33	82.5	16	1	US-08-475-041-11	Sequence 11, Appl
63	33	82.5	16	1	US-08-484-773-11	Sequence 11, Appl
64	33	82.5	16	1	US-08-361-864-15	Sequence 15, Appl
65	33	82.5	17	1	US-08-753-781-23	Sequence 23, Appl
66	33	82.5	19	1	US-08-335-832-11	Sequence 11, Appl
67	33	82.5	19	2	US-09-141-127-5	Sequence 5, Appl
68	33	82.5	20	1	US-08-753-781-5	Sequence 5, Appl
69	33	82.5	53	2	US-09-621-976-5473	Sequence 5473, Ap
70	33	82.5	119	2	US-09-902-540-10158	Sequence 10158, A
71	33	82.5	119	2	US-09-830-954A-28	Sequence 28, Appl
72	33	82.5	131	2	US-10-104-047-3039	Sequence 3039, Ap
73	33	82.5	172	2	US-09-252-991A-19020	Sequence 19020, A
74	33	82.5	214	2	US-09-902-540-14768	Sequence 14768, A
75	33	82.5	217	2	US-09-949-016-8819	Sequence 8819, Ap
76	33	82.5	217	2	US-09-949-016-8820	Sequence 8820, Ap
77	33	82.5	217	2	US-09-949-016-8821	Sequence 8821, Ap
78	33	82.5	221	2	US-09-252-991A-26636	Sequence 26636, A
79	33	82.5	239	2	US-09-830-954A-6	Sequence 6, Appl
80	33	82.5	264	2	US-10-114-716A-46	Sequence 46, Appl
81	33	82.5	286	2	US-09-252-991A-25755	Sequence 25755, A
82	33	82.5	326	2	US-09-252-991A-21574	Sequence 21574, A
83	33	82.5	361	2	US-09-538-092-1331	Sequence 1331, Ap
84	33	82.5	409	2	US-09-248-796A-15496	Sequence 15496, A
85	33	82.5	416	2	US-09-902-540-14976	Sequence 14976, A
86	33	82.5	427	2	US-09-252-991A-32373	Sequence 32373, A
87	33	82.5	464	2	US-09-252-991A-20795	Sequence 20795, A
88	33	82.5	528	2	US-09-490-291-8	Sequence 8, Appl
89	33	82.5	607	2	US-09-252-991A-26647	Sequence 26647, A
90	33	82.5	637	2	US-09-949-016-8152	Sequence 8152, Ap
91	33	82.5	643	2	US-09-538-092-844	Sequence 844, App
92	33	82.5	687	2	US-09-270-767-46104	Sequence 46104, A
93	33	82.5	699	2	US-09-949-016-6652	Sequence 6652, Ap
94	33	82.5	699	2	US-10-144-198-40	Sequence 40, Appl
95	33	82.5	710	2	US-09-949-016-6808	Sequence 6808, Ap
96	33	82.5	710	2	US-10-144-198-39	Sequence 39, Appl
97	33	82.5	714	2	US-09-828-303-21	Sequence 21, Appl
98	33	82.5	736	2	US-09-252-991A-31194	Sequence 31194, A
99	33	82.5	736	2	US-09-949-016-6807	Sequence 6807, Ap
100	33	82.5	736	2	US-10-144-198-38	Sequence 38, Appl

ALIGNMENTS

RESULT 1
US-09-252-991A-20941
; Sequence 20941, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20941
; LENGTH: 449
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20941

Query Match 100.0%; Score 40; DB 2; Length 449;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
|||:||||
Db 329 GGIGDGG 335

RESULT 2
US-09-248-796A-24658
; Sequence 24658, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 24658
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-24658

Query Match 97.5%; Score 39; DB 2; Length 200;
Best Local Similarity 85.7%; Pred. No. 56;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
|||:||||
Db 178 GGIGDGG 184

RESULT 3
US-09-902-540-15733
; Sequence 15733, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.

; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 15733
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-15733

Query Match 97.5%; Score 39; DB 2; Length 233;
Best Local Similarity 85.7%; Pred. No. 65;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
|||:||||
Db 156 GGIGDGG 162

RESULT 4
US-09-252-991A-32661
; Sequence 32661, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32661
; LENGTH: 430
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32661

Query Match 95.0%; Score 38; DB 2; Length 430;
Best Local Similarity 85.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
|||:||||
Db 64 GGIGDGG 70

RESULT 5
US-09-489-039A-8382
; Sequence 8382, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 8382
; LENGTH: 142
; TYPE: PRT

; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8382

Query Match 87.5%; Score 35; DB 2; Length 142;
Best Local Similarity 85.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGIGDGG 7
|||
Db 95 GGAGDGG 101

RESULT 6

US-09-252-991A-29526
; Sequence 29526, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 29526

; LENGTH: 281

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-29526

Query Match 87.5%; Score 35; DB 2; Length 281;
Best Local Similarity 85.7%; Pred. No. 3.1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGIGDGG 7
|||
Db 26 GGAGDGG 32

RESULT 7

US-09-270-767-45414
; Sequence 45414, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:

; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

; FILE REFERENCE: File Reference: 7326-094

; CURRENT APPLICATION NUMBER: US/09/270,767

; CURRENT FILING DATE: 1999-03-17

; NUMBER OF SEQ ID NOS: 62517

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 45414

; LENGTH: 327

; TYPE: PRT

; ORGANISM: Drosophila melanogaster

; FEATURE:

; OTHER INFORMATION: Xaa means any amino acid

US-09-270-767-45414

Query Match 87.5%; Score 35; DB 2; Length 327;
Best Local Similarity 85.7%; Pred. No. 3.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGIGDGG 7
|||
Db 99 GGAGDGG 105

RESULT 8

US-09-489-039A-7341
; Sequence 7341, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA

; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

; CURRENT APPLICATION NUMBER: US/09/489,039A

; PRIOR FILING DATE: 2000-01-27

; PRIOR APPLICATION NUMBER: US 60/117,747

US-09-252-991A-27369
; Sequence 27369, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 27369

; LENGTH: 334

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-27369

Query Match 87.5%; Score 35; DB 2; Length 334;
Best Local Similarity 71.4%; Pred. No. 3.7e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGIGDGG 7
|||
Db 237 GGVGEGG 243

RESULT 9

US-09-902-540-16071
; Sequence 16071, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:

; APPLICANT: Goldman, Barry S.

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Wiegand, Roger C.

; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof

; FILE REFERENCE: 38-10(15849)H

; CURRENT APPLICATION NUMBER: US/09/902,540

; CURRENT FILING DATE: 2001-07-10

; PRIOR APPLICATION NUMBER: 60/217,883

; PRIOR FILING DATE: 2000-07-10

; NUMBER OF SEQ ID NOS: 16825

; SEQ ID NO 16071

; LENGTH: 341

; TYPE: PRT

; ORGANISM: Myxococcus xanthus

US-09-902-540-16071

Query Match 87.5%; Score 35; DB 2; Length 341;
Best Local Similarity 85.7%; Pred. No. 3.8e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGIGDGG 7
|||
Db 46 GGIGNGG 52

RESULT 10

US-09-489-039A-7341
; Sequence 7341, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA

; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

; CURRENT APPLICATION NUMBER: US/09/489,039A

; PRIOR FILING DATE: 2000-01-27

; PRIOR APPLICATION NUMBER: US 60/117,747

; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 7341
; LENGTH: 405
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7341

Query Match 87.5%; Score 35; DB 2; Length 405;
Best Local Similarity 85.7%; Pred. No. 4.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
|||:||||
Db 47 GGAGDGG 53

RESULT 11
US-09-252-991A-24561
; Sequence 24561, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24561
; LENGTH: 406
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24561

Query Match 87.5%; Score 35; DB 2; Length 406;
Best Local Similarity 85.7%; Pred. No. 4.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
|||:||||
Db 342 GGAGDGG 348

RESULT 12
US-09-252-991A-22415
; Sequence 22415, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22415
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22415

Query Match 87.5%; Score 35; DB 2; Length 416;
Best Local Similarity 71.4%; Pred. No. 4.6e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
|||:||||
Db 130 GGVGEGG 136

RESULT 13
US-09-252-991A-23585
; Sequence 23585, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23585
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23585

Query Match 87.5%; Score 35; DB 2; Length 416;
Best Local Similarity 85.7%; Pred. No. 4.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
|||:||||
Db 254 GGAGDGG 260

RESULT 14
US-09-252-991A-31062
; Sequence 31062, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31062
; LENGTH: 535
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31062

Query Match 87.5%; Score 35; DB 2; Length 535;
Best Local Similarity 71.4%; Pred. No. 5.8e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
|||:||||
Db 503 GGVGEGG 509

RESULT 15
US-08-484-905-4
; Sequence 4, Application US/08484905
; Patent No. 5976551
; GENERAL INFORMATION:


```

; APPLICANT: Mottez, Estelle
; APPLICANT: Abastado, Jean-Pierre
; APPLICANT: Kourilsky, Philippe
; TITLE OF INVENTION: An Altered Major Histocompatibility
; DETERMINANT AND METHODS FOR USING THE
; DETERMINANT
; NUMBER OF SEQUENCES: 127
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADRESSEE: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US 07/801,818
; FILING DATE: 05-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/792,473
; FILING DATE: 15-NOV-1991
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Potter, Jane E. R.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 03495.0106-03000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4400
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-484-905-4

Query Match 85.0%; Score 34; DB 1; Length 10;
Best Local Similarity 85.7%; Pred. No. 18;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGIGDGG 7
Db 1 GGIGSGG 7

RESULT 16
US-08-481-985B-4
; Sequence 4, Application US/08481985B
; Patent No. 601146
; GENERAL INFORMATION:
; APPLICANT: Mottez, Estelle
; APPLICANT: Abastado, Jean-Pierre
; APPLICANT: Kourilsky, Philippe
; TITLE OF INVENTION: Altered Major Histocompatibility Complex
; TITLE OF INVENTION:
; NUMBER OF SEQUENCES: 148
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADRESSEE: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; ZIP: 20005-3315

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US 08/481,985B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/801,818
; FILING DATE: 05-DEC-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/792,473
; FILING DATE: 15-NOV-1991
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0106-04000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-481-985B-4

Query Match 85.0%; Score 34; DB 2; Length 10;
Best Local Similarity 85.7%; Pred. No. 18;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGIGDGG 7
Db 1 GGIGSGG 7

RESULT 17
US-08-370-476-4
; Sequence 4, Application US/08370476
; Patent No. 6153408
; GENERAL INFORMATION:
; APPLICANT: Mottez, Estelle
; APPLICANT: Abastado, Jean-Pierre
; APPLICANT: Kourilsky, Philippe
; APPLICANT: Lone, Yu-Chun
; APPLICANT: Ojcius, David
; APPLICANT: Casrouge, Armanda
; TITLE OF INVENTION: Altered Major Histocompatibility Complex
; TITLE OF INVENTION:
; NUMBER OF SEQUENCES: 127
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADRESSEE: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US 08/370,476
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/117,575

```

; FILING DATE: 07-SEP-1993
; APPLICATION NUMBER: US 08/072,787
; FILING DATE: 06-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/801,818
; FILING DATE: 05-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/792,473
; FILING DATE: 15-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 05243.0001-01000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-370-476-4

Query Match 85.0%; Score 34; DB 2; Length 10;
Best Local Similarity 85.7%; Pred. No. 18;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
||| |||
Db 1 GGIGSGG 7

RESULT 18
US-08-484-905-6
; Sequence 6, Application US/08484905
; Patent No. 5976551
; GENERAL INFORMATION:
; APPLICANT: Mottez, Estelle
; APPLICANT: Abastado, Jean-Pierre
; APPLICANT: Kourilsky, Philippe
; TITLE OF INVENTION: An Altered Major Histocompatibility
; TITLE OF INVENTION: Complex(MHC) Determinant and Methods for Using the
; TITLE OF INVENTION: Determinant
; NUMBER OF SEQUENCES: 127
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS-/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,905
; FILING DATE: 07-JUNE-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/792,473
; FILING DATE: 15-NOV-1991
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Potter, Jane E. R.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 03495.0106-03000

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-484-905-6

Query Match 85.0%; Score 34; DB 1; Length 13;
Best Local Similarity 85.7%; Pred. No. 24;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
||| |||
Db 1 GGIGSGG 7

RESULT 19
US-08-481-985B-6
; Sequence 6, Application US/08481985B
; Patent No. 6011146
; GENERAL INFORMATION:
; APPLICANT: Mottez, Estelle
; APPLICANT: Abastado, Jean-Pierre
; APPLICANT: Kourilsky, Philippe
; TITLE OF INVENTION: Altered Major Histocompatibility Complex
; TITLE OF INVENTION:
; NUMBER OF SEQUENCES: 148
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,985B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/801,818
; FILING DATE: 05-DEC-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/792,473
; FILING DATE: 15-NOV-1991
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0106-04000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-481-985B-6

Query Match 85.0%; Score 34; DB 2; Length 13;
Best Local Similarity 85.7%; Pred. No. 24;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
|||||
Db 1 GGIGSGG 7

RESULT 20

US-08-370-476-6
; Sequence 6, Application US/08370476
; Patent No. 6153408
; GENERAL INFORMATION:
; APPLICANT: Mottez, Estelle
; APPLICANT: Abastado, Jean-Pierre
; APPLICANT: Kourilsky, Philippe
; APPLICANT: Lorne, Yu-Chun
; APPLICANT: Ojcius, David
; APPLICANT: Casrouge, Amanda
; TITLE OF INVENTION: Altered Major Histocompatibility Complex
; TITLE OF INVENTION:
; NUMBER OF SEQUENCES: 127
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,476
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/117,575
; FILING DATE: 07-SEP-1993
; APPLICATION NUMBER: US 08/072,787
; FILING DATE: 06-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/801,818
; FILING DATE: 05-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/792,473
; FILING DATE: 15-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 05243.0001-01000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-370-476-6

Query Match 85.0%; Score 34; DB 2; Length 13;
Best Local Similarity 85.7%; Pred. No. 24;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
|||||
Db 1 GGIGSGG 7

RESULT 21

US-08-484-905-8

; Sequence 8, Application US/08484905
; Patent No. 5976551
; GENERAL INFORMATION:
; APPLICANT: Mottez, Estelle
; APPLICANT: Abastado, Jean-Pierre
; APPLICANT: Kourilsky, Philippe
; TITLE OF INVENTION: An Altered Major Histocompatibility
; TITLE OF INVENTION: Complex(MHC) Determinant and Methods for Using the
; TITLE OF INVENTION: Determinant
; NUMBER OF SEQUENCES: 127
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS-/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,905
; FILING DATE: 07-JUNE-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/801,818
; FILING DATE: 05-DEC-1991
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/792,473
; FILING DATE: 15-NOV-1991
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Potter, Jane E. R.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 03495.0106-03000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-484-905-8

Query Match 85.0%; Score 34; DB 1; Length 15;
Best Local Similarity 85.7%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
|||||
Db 1 GGIGSGG 7

RESULT 22

US-08-481-985B-8
; Sequence 8, Application US/08481985B
; Patent No. 6011146
; GENERAL INFORMATION:
; APPLICANT: Mottez, Estelle
; APPLICANT: Abastado, Jean-Pierre
; APPLICANT: Kourilsky, Philippe
; TITLE OF INVENTION: Altered Major Histocompatibility Complex
; TITLE OF INVENTION:
; NUMBER OF SEQUENCES: 148
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner

STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481,985B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/801,818
FILING DATE: 05-DEC-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/792,473
FILING DATE: 15-NOV-1991
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03495.0106-04000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-481-985B-8

Query Match 85.0%; Score 34; DB 2; Length 15;
Best Local Similarity 85.7%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
Db 1 GGIGSGG 7

RESULT 23
US-08-370-476-8
Sequence 8, Application US/08370476
Patent No. 6153408
GENERAL INFORMATION:
APPLICANT: Mottez, Estelle
APPLICANT: Abastado, Jean-Pierre
APPLICANT: Kourilsky, Philippe
APPLICANT: Lone, Yu-Chun
APPLICANT: Ojcius, David
APPLICANT: Casrouge, Armanda
TITLE OF INVENTION: Altered Major Histocompatibility Complex
NUMBER OF SEQUENCES: 127
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/370,476
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/117,575
FILING DATE: 07-SEP-1993
APPLICATION NUMBER: US 08/072,787
FILING DATE: 06-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/801,818
FILING DATE: 05-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/792,473
FILING DATE: 15-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 05243.0001-01000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-370-476-8

Query Match 85.0%; Score 34; DB 2; Length 15;
Best Local Similarity 85.7%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
Db 1 GGIGSGG 7

RESULT 24
US-08-484-905-12
Sequence 12, Application US/08484905
Patent No. 5976551
GENERAL INFORMATION:
APPLICANT: Mottez, Estelle
APPLICANT: Abastado, Jean-Pierre
APPLICANT: Kourilsky, Philippe
TITLE OF INVENTION: An Altered Major Histocompatibility
TITLE OF INVENTION: Complex(MHC) Determinant and Methods for Using the
TITLE OF INVENTION: Determinant
NUMBER OF SEQUENCES: 127
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,905
FILING DATE: 07-JUNE-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/801,818
FILING DATE: 05-DEC-1991
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/792,473

;
; FILING DATE: 15-NOV-1991
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Potter, Jane E. R.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 03495.0106-03000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-484-905-12

Query Match 85.0%; Score 34; DB 1; Length 19;
Best Local Similarity 85.7%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGIGDGG 7
 |||||
Db 1 GGIGSGG 7

RESULT 25
US-08-481-985B-12
; Sequence 12, Application US/08481985B
; Patent No. 6011146
; GENERAL INFORMATION:
; APPLICANT: Mottez, Estelle
; APPLICANT: Abastado, Jean-Pierre
; APPLICANT: Kourilsky, Philippe
; TITLE OF INVENTION: Altered Major Histocompatibility Complex
; NUMBER OF SEQUENCES: 148
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,985B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/801,818
; FILING DATE: 05-DEC-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/792,473
; FILING DATE: 15-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0106-04000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

;
; MOLECULE TYPE: peptide
US-08-481-985B-12

Query Match 85.0%; Score 34; DB 2; Length 19;
Best Local Similarity 85.7%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGIGDGG 7
 |||||
Db 1 GGIGSGG 7

RESULT 26
US-08-370-476-12
; Sequence 12, Application US/08370476
; Patent No. 6153408
; GENERAL INFORMATION:
; APPLICANT: Mottez, Estelle
; APPLICANT: Abastado, Jean-Pierre
; APPLICANT: Kourilsky, Philippe
; APPLICANT: Lone, Yu-Chun
; APPLICANT: Ojcius, David
; APPLICANT: Casrouge, Armanda
; TITLE OF INVENTION: Altered Major Histocompatibility Complex
; NUMBER OF SEQUENCES: 127
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,476
; FILING DATE: 06-JUN-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/117,575
; FILING DATE: 07-SEP-1993
; APPLICATION NUMBER: US 08/072,787
; FILING DATE: 06-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/801,818
; FILING DATE: 05-DEC-1991
; APPLICATION NUMBER: US 07/792,473
; FILING DATE: 15-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 05243.0001-01000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-370-476-12

Query Match 85.0%; Score 34; DB 2; Length 19;
Best Local Similarity 85.7%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGIGDGG 7

```
Db      1 GGIGSGG 7
;
; TITLE OF INVENTION:
; NUMBER OF SEQUENCES: 148
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; TITLE OF INVENTION: An Altered Major Histocompatibility
; DETERMINANT AND METHODS FOR USING THE
; DETERMINANT
; NUMBER OF SEQUENCES: 127
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,905
; FILING DATE: 07-JUNE-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/801,818
; FILING DATE: 05-DEC-1991
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Porter, Jane E. R.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 03495.0106-03000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4400
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-484-905-14

Query Match      85.0%; Score 34; DB 1; Length 21;
Best Local Similarity 85.7%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GGIGDGG 7
      |||||
Db      1 GGIGSGG 7

RESULT 29
US-08-370-476-14
; Sequence 14, Application US/08370476
; Patent No. 6153408
; GENERAL INFORMATION:
; APPLICANT: Mottez, Estelle
; APPLICANT: Abastado, Jean-Pierre
; APPLICANT: Kourilsky, Philippe
; APPLICANT: Lone, Yu-Chun
; APPLICANT: Ojcius, David
; APPLICANT: Casrouge, Armanda
; TITLE OF INVENTION: Altered Major Histocompatibility Complex
; NUMBER OF SEQUENCES: 127
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

Db      1 GGIGSGG 7
;
; TITLE OF INVENTION:
; NUMBER OF SEQUENCES: 148
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; TITLE OF INVENTION: An Altered Major Histocompatibility
; DETERMINANT AND METHODS FOR USING THE
; DETERMINANT
; NUMBER OF SEQUENCES: 127
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,905
; FILING DATE: 07-JUNE-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/801,818
; FILING DATE: 05-DEC-1991
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Porter, Jane E. R.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 03495.0106-03000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4400
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-484-905-14

Query Match      85.0%; Score 34; DB 1; Length 21;
Best Local Similarity 85.7%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GGIGDGG 7
      |||||
Db      1 GGIGSGG 7

RESULT 28
US-08-481-985B-14
; Sequence 14, Application US/08481985B
; Patent No. 6011146
; GENERAL INFORMATION:
; APPLICANT: Mottez, Estelle
; APPLICANT: Abastado, Jean-Pierre
; APPLICANT: Kourilsky, Philippe
; TITLE OF INVENTION: Altered Major Histocompatibility Complex
```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/370,476
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/117,575
FILING DATE: 07-SEP-1993
APPLICATION NUMBER: US 08/072,787
FILING DATE: 06-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/801,818
FILING DATE: 05-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/792,473
FILING DATE: 15-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 05243.0001-01000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-370-476-14

Query Match 85.0%; Score 34; DB 2; Length 21;
Best Local Similarity 85.7%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGIGDGG 7
| | | | |
Db 1 GGIGSGG 7

RESULT 30
US-08-461-597-3
; Sequence 3, Application US/08461597
; Patent No. 5759834
; GENERAL INFORMATION:
; APPLICANT: Atassi, M. Zouhair
; TITLE OF INVENTION: Synthetic Sterically-Constrained
; TITLE OF INVENTION: Catalysts
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Thomas D. Paul;
; STREET: 1301 McKinney, Suite 5100
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77010-3095
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,597
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/063,640
; FILING DATE: 18-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul, Thomas D.
; REGISTRATION NUMBER: 32,714

REFERENCE/DOCKET NUMBER: D-5730
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713/651-5325
TELEFAX: 713/651-4246
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
TOPOLOGY: circular
MOLECULE TYPE: protein
HYPOTHETICAL: NO
FRAGMENT TYPE: not applicable
US-08-461-597-3

Query Match 85.0%; Score 34; DB 1; Length 37;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGIGDG 6
| | | | |
Db 18 GGIGDG 23

RESULT 31
US-08-535-298-3
; Sequence 3, Application US/08535298
; Patent No. 5861477
; GENERAL INFORMATION:
; APPLICANT: Atassi, Zouhair
; TITLE OF INVENTION: Synthetic Sterically-Constrained
; TITLE OF INVENTION: Catalysts
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: C. Steven McDaniel;
; STREET: 600 Travis Street, Suite 1850
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77002-2912
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Microsoft Word 97 (running under MS Windows 95)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/535,298
; FILING DATE: 19-JUNE-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: C. Steven McDaniel
; REGISTRATION NUMBER: 33,962
; REFERENCE/DOCKET NUMBER: 1515-00103
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713/238-8000
; TELEFAX: 713/238-8008
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37 amino acids
; TYPE: amino acid
; TOPOLOGY: circular
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FRAGMENT TYPE: not applicable
US-08-535-298-3

Query Match 85.0%; Score 34; DB 1; Length 37;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGIGDG 6
| | | | |
Db 18 GGIGDG 23

RESULT 32
PCT-US94-05569A-3
; Sequence 3, Application PC/TUS9405569A

; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Synthetic Sterically-Constrained
; TITLE OF INVENTION: Catalysts
; NUMBER OF SEQUENCES: 6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/05569A
; FILING DATE: 18-MAY-1994

; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/063,640
; FILING DATE:

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 37 amino acids

; TYPE: amino acid

; TOPOLOGY: circular

; MOLECULE TYPE: protein

; HYPOTHETICAL: NO

; FRAGMENT TYPE: not applicable

PCT-US94-05569A-3

Query Match 85.0%; Score 34; DB 4; Length 37;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDG 6
|||
Db 18 GGIGDG 23

RESULT 33
PCT-US94-05569-3
; Sequence 3, Application PC/TUS9405569

; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Synthetic Sterically-Constrained Catalysts

; NUMBER OF SEQUENCES: 6

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US94/05569

; FILING DATE: 18-MAY-1994

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/063,640

; FILING DATE: 18-MAY-1993

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 37 amino acids

; TYPE: amino acid

; TOPOLOGY: circular

; MOLECULE TYPE: protein

; HYPOTHETICAL: NO

; FRAGMENT TYPE: not applicable

PCT-US94-05569-3

Query Match 85.0%; Score 34; DB 4; Length 37;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDG 6
|||
Db 18 GGIGDG 23

RESULT 34

US-09-248-796A-21808

; Sequence 21808, Application US/09248796A

; Patent No. 6747137

; GENERAL INFORMATION:

; APPLICANT: Keith Weinstock et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS

; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.132

; CURRENT APPLICATION NUMBER: US/09/248,796A

; CURRENT FILING DATE: 1999-02-12

; PRIOR APPLICATION NUMBER: US 60/074,725

; PRIOR FILING DATE: 1998-02-13

; PRIOR APPLICATION NUMBER: US 60/096,409

; PRIOR FILING DATE: 1998-08-13

; NUMBER OF SEQ ID NOS: 28208

; SEQ ID NO 21808

; LENGTH: 170

; TYPE: PRT

; ORGANISM: Candida albicans

; FEATURE:

; NAME/KEY: UNSURE

; LOCATION: (12)

; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknown

US-09-248-796A-21808

Query Match 85.0%; Score 34; DB 2; Length 170;
Best Local Similarity 71.4%; Pred. No. 2.8e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
|||
Db 20 GGVGNGG 26

RESULT 35

US-09-270-767-42925

; Sequence 42925, Application US/09270767

; Patent No. 6703491

; GENERAL INFORMATION:

; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

; FILE REFERENCE: File Reference: 7326-094

; CURRENT APPLICATION NUMBER: US/09/270,767

; CURRENT FILING DATE: 1999-03-17

; NUMBER OF SEQ ID NOS: 62517

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 42925

; LENGTH: 256

; TYPE: PRT

; ORGANISM: Drosophila melanogaster

US-09-270-767-42925

Query Match 85.0%; Score 34; DB 2; Length 256;
Best Local Similarity 85.7%; Pred. No. 4.1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
|||
Db 133 GGSDDGG 139

RESULT 36

US-09-498-520A-12

; Sequence 12, Application US/09498520A

; Patent No. 6613553

; GENERAL INFORMATION:

; APPLICANT: Rock, Charles O


```
; APPLICANT: Heath, Richard J
; TITLE OF INVENTION: No. 6613553el Enoyl Reductases and Methods of Use Thereof
; FILE REFERENCE: SJ-0022
; CURRENT APPLICATION NUMBER: US/09/498,520A
; CURRENT FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Clostridium acetobutylicum
US-09-498-520A-12

Query Match      85.0%; Score 34; DB 2; Length 310;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GGIGDG 6
Db      170 GGIGDG 175

RESULT 37
US-09-498-520A-10
; Sequence 10, Application US/09498520A
; Patent No. 6613553
; GENERAL INFORMATION:
; APPLICANT: Heath, Richard J
; TITLE OF INVENTION: No. 6613553el Enoyl Reductases and Methods of Use Thereof
; FILE REFERENCE: SJ-0022
; CURRENT APPLICATION NUMBER: US/09/498,520A
; CURRENT FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-498-520A-10

Query Match      85.0%; Score 34; DB 2; Length 332;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GGIGDG 6
Db      183 GGIGDG 188

RESULT 38
US-09-134-000C-4988
; Sequence 4988, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4988
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-4988

Query Match      85.0%; Score 34; DB 2; Length 332;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
```

```
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GGIGDG 6
Db      183 GGIGDG 188

RESULT 39
US-09-252-991A-27327
; Sequence 27327, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27327
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27327

Query Match      85.0%; Score 34; DB 2; Length 341;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GGIGDG 6
Db      313 GGIGDG 318

RESULT 40
US-09-252-991A-25656
; Sequence 25656, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25656
; LENGTH: 356
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25656

Query Match      85.0%; Score 34; DB 2; Length 356;
Best Local Similarity 71.4%; Pred. No. 5.6e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GGIGDGG 7
Db      62 GGLGEGG 68

RESULT 41
US-09-149-476-751
; Sequence 751, Application US/09149476
; Patent No. 6420526
```

GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 186 Human Secreted proteins
FILE REFERENCE: P2002P1
CURRENT APPLICATION NUMBER: US/09/149,476
CURRENT FILING DATE: 1998-09-08
EARLIER APPLICATION NUMBER: PCT/US98/04493
EARLIER FILING DATE: 1998-03-06
EARLIER APPLICATION NUMBER: 60/040,162
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,333
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/038,621
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,626
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,334
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,336
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,163
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/047,600
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,615
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,597
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,502
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,633
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,583
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,617
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,618
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,503
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,592
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,581
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,584
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,500
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,587
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,492
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,598
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,613
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,582
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,596
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,612
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,632
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,601
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,580
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,568
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,314
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,569
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,311
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,671
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,674
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,669
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,312
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,313
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,672
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,315
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/056,886
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,877
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,889
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,893
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,662
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,872
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,882
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,637
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,903
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,888
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,879
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,880
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,894
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,911
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,636
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,874
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,910
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,864
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,631
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,845
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,892
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,761
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,595
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,599
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,588
EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,585
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,586
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,590
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,594
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,589
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,593
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,614
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,578
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,576
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,501
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,670
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/056,632
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,664
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,876
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,881
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,909
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,875
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,669
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060
; EARLIER FILING DATE: 1997-10-02

Query Match 85.0%; Score 34; DB 2; Length 413;
Best Local Similarity 85.7%; Pred. No. 6.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGIGDGG 7
| | | | |
Db 195 GNGDGG 201

RESULT 42
US-09-252-991A-29256
; Sequence 29256, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29256
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29256

Query Match 85.0%; Score 34; DB 2; Length 434;
Best Local Similarity 71.4%; Pred. No. 6.8e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGIGDGG 7
| | | | |
Db 254 GGLGEGG 260

RESULT 43
US-09-252-991A-26584
; Sequence 26584, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26584
; LENGTH: 469
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26584

Query Match 85.0%; Score 34; DB 2; Length 469;
Best Local Similarity 71.4%; Pred. No. 7.3e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGIGDGG 7
| | | | |
Db 343 GGLGEGG 349

RESULT 44
US-09-489-039A-9961
; Sequence 9961, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9961
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9961

Query Match 85.0%; Score 34; DB 2; Length 614;
Best Local Similarity 100.0%; Pred. No. 9.5e+02;

```
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GIGDGG 7
Db 95 GIGDGG 100

RESULT 45
US-09-949-016-7243
; Sequence 7243, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7243
; LENGTH: 693
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7243

Query Match 85.0%; Score 34; DB 2; Length 693;
Best Local Similarity 85.7%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
Db 604 GNIGDGG 610

RESULT 46
US-08-999-774A-6
; Sequence 6, Application US/08999774A
; Patent No. 6274312
; GENERAL INFORMATION:
; APPLICANT: Gish, Kurt C.
; APPLICANT: Seghezzi, Wolfgang
; APPLICANT: Shanahan, Frances
; APPLICANT: Lees, Emma M.
; APPLICANT: McClanahan, Terrill K.
; TITLE OF INVENTION: Intracellular Regulatory Molecules;
; TITLE OF INVENTION: Related Reagents
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; APPLICATION NUMBER: US/08/999,774A
; FILING DATE: 10-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/032,818
; FILING DATE: 11-DEC-1996
```

```
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34,090
; REFERENCE/DOCKET NUMBER: DX0646
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650)852-9196
; TELEFAX: (650)496-1200
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 806 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-999-774A-6

Query Match 85.0%; Score 34; DB 2; Length 806;
Best Local Similarity 85.7%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
Db 588 GNIGDGG 594

RESULT 47
US-09-949-016-6786
; Sequence 6786, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6786
; LENGTH: 806
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6786

Query Match 85.0%; Score 34; DB 2; Length 806;
Best Local Similarity 85.7%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
Db 588 GNIGDGG 594

RESULT 48
US-09-252-991A-24163
; Sequence 24163, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
```

; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24163
; LENGTH: 912
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24163

Query Match 85.0%; Score 34; DB 2; Length 912;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GIGDGG 7
|||||
Db 211 GIGDGG 216

RESULT 49

US-09-809-920-2

; Sequence 2, Application US/09809920

; Patent No. 6812326

; GENERAL INFORMATION:

; APPLICANT: Sato, Takaaki

; TITLE OF INVENTION: TREX, A NOVEL GENE OF TRAF-INTERACTING
; EXT GENE FAMILY AND DIAGNOSTIC AND THERAPEUTIC USES
; THEREOF

; NUMBER OF SEQUENCES: 37

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Cooper & Dunham LLP

; STREET: 1185 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/809,920

; FILING DATE: 16-Mar-2001

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/09/156,191

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: White, John P.

; REGISTRATION NUMBER: 28,678

; REFERENCE/DOCKET NUMBER: 0575/51902

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 278-0400

; TELEFAX: (212) 391-0525

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 918 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Query Match

85.0%; Score 34; DB 2; Length 918;

Best Local Similarity 71.4%; Pred. No. 1.4e+03;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GIGDGG 7

|||||

Db 10 GGVGGG 16

RESULT 50

US-08-482-880-10

; Sequence 10, Application US/08482880

; Patent No. 5736122
; GENERAL INFORMATION:
; APPLICANT: Dean, Richard T
; APPLICANT: Lister-James, John
; TITLE OF INVENTION: Technetium-99m Labeled Peptides for
; TITLE OF INVENTION: Thrombus Imaging
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Allegretti, Ltd.
; STREET: 10 South Wacker Drive Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,880
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5736122nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 92,216-L
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-482-880-10

Query Match 82.5%; Score 33; DB 1; Length 11;
Best Local Similarity 85.7%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGIGDGG 7

|||||

Db 1 GGRGDGG 7

Search completed: February 27, 2006, 09:51:27
Job time : 48 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: February 27, 2006, 10:01:20 ; Search time 164 Seconds
(without alignments)
17.834 Million cell updates/sec

Title: US-10-027-015B-2

Perfect score: 40

Sequence: 1 GGIGDG 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : Published Applications AA_Main:*

1: /cgn2_6/ptodata/1/pubpaa/us07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/us08_PUBCOMB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/us09_PUBCOMB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/us10a_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/us10b_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/us11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	7	5	US-10-027-015A-2
2	40	100.0	1090	4	US-10-149-310-176
3	39	97.5	60	4	US-10-437-963-103936
4	39	97.5	88	5	US-10-732-923-3741
5	39	97.5	156	4	US-10-437-963-170096
6	39	97.5	157	3	US-09-864-761-44299
7	39	97.5	184	4	US-10-437-963-149979
8	39	97.5	250	3	US-09-925-298-517
9	39	97.5	250	4	US-10-102-806-517
10	39	97.5	333	4	US-10-437-963-135984
11	39	97.5	377	4	US-10-425-115-277369
12	39	97.5	399	5	US-10-450-763-50000
13	39	97.5	897	4	US-10-437-963-183531
14	38	95.0	37	4	US-10-029-386-29024
15	38	95.0	120	4	US-10-437-963-180674
16	38	95.0	122	4	US-10-425-115-201529
17	38	95.0	170	4	US-10-437-963-154584
18	38	95.0	181	4	US-10-437-963-141956
19	38	95.0	214	4	US-10-437-963-111588
20	38	95.0	300	4	US-10-437-963-157722
21	38	95.0	427	6	US-11-097-143-9876
22	38	95.0	1152	4	US-10-437-963-173239
23	37	92.5	146	4	US-10-437-963-121618
24	36	90.0	93	4	US-10-425-115-229793
25	36	90.0	117	4	US-10-425-115-205711
26	36	90.0	170	4	US-10-437-963-196524
27	36	90.0	187	4	US-10-437-963-130467

28	36	90.0	290	4	US-10-437-963-159222	Sequence 159222,
29	36	90.0	292	4	US-10-425-115-211110	Sequence 211110,
30	36	90.0	383	4	US-10-437-963-149137	Sequence 149137,
31	36	90.0	405	4	US-10-437-963-172290	Sequence 172290,
32	36	90.0	456	4	US-10-156-761-10244	Sequence 10244, A
33	36	90.0	813	4	US-10-437-963-159271	Sequence 159271,
34	36	90.0	1278	4	US-10-156-761-10231	Sequence 10231, A
35	36	90.0	11088	4	US-10-203-295-7	Sequence 7, Appl
36	36	90.0	11096	5	US-10-732-923-20557	Sequence 20557, A
37	35	87.5	61	3	US-09-864-761-36444	Sequence 36444, A
38	35	87.5	61	3	US-09-864-761-42760	Sequence 42760, A
39	35	87.5	70	4	US-10-424-599-192048	Sequence 192048,
40	35	87.5	75	4	US-10-437-963-109585	Sequence 109585,
41	35	87.5	84	4	US-10-437-963-142504	Sequence 142504,
42	35	87.5	86	4	US-10-424-599-184457	Sequence 184457,
43	35	87.5	92	4	US-10-425-115-193758	Sequence 193758,
44	35	87.5	96	4	US-10-437-963-153974	Sequence 153974,
45	35	87.5	97	4	US-10-437-963-129822	Sequence 129822,
46	35	87.5	100	4	US-10-767-701-42596	Sequence 42596, A
47	35	87.5	108	4	US-10-425-115-189769	Sequence 189769,
48	35	87.5	119	4	US-10-424-599-200912	Sequence 200912,
49	35	87.5	128	4	US-10-437-963-140184	Sequence 140184,
50	35	87.5	132	4	US-10-425-115-188673	Sequence 188673,
51	35	87.5	136	4	US-10-767-701-56761	Sequence 56761, A
52	35	87.5	137	3	US-09-864-408A-4054	Sequence 4054, Ap
53	35	87.5	143	4	US-10-425-115-201325	Sequence 201325,
54	35	87.5	157	4	US-10-437-963-142505	Sequence 142505,
55	35	87.5	165	4	US-10-374-780A-505	Sequence 505, App
56	35	87.5	170	4	US-10-425-115-125034	Sequence 225034,
57	35	87.5	193	4	US-10-437-963-103684	Sequence 103684,
58	35	87.5	206	4	US-10-437-963-145746	Sequence 145746,
59	35	87.5	220	4	US-10-437-963-192824	Sequence 192824,
60	35	87.5	223	4	US-10-424-599-244310	Sequence 244310,
61	35	87.5	227	4	US-10-437-963-137129	Sequence 137129,
62	35	87.5	228	5	US-10-732-923-5699	Sequence 5699, Ap
63	35	87.5	236	4	US-10-437-963-203317	Sequence 203317,
64	35	87.5	237	4	US-10-437-963-151884	Sequence 151884,
65	35	87.5	248	5	US-10-739-930-10469	Sequence 10469, A
66	35	87.5	270	4	US-10-767-701-34834	Sequence 34834, A
67	35	87.5	277	4	US-10-424-599-192146	Sequence 192146,
68	35	87.5	281	5	US-10-450-763-43699	Sequence 43699, A
69	35	87.5	290	4	US-10-437-963-157718	Sequence 157718,
70	35	87.5	325	4	US-10-334-143-56	Sequence 56, Appl
71	35	87.5	334	6	US-11-097-143-28095	Sequence 28095, A
72	35	87.5	350	4	US-10-437-963-112926	Sequence 112926,
73	35	87.5	357	3	US-09-864-761-35807	Sequence 35807, A
74	35	87.5	360	5	US-10-739-930-8418	Sequence 8418, Ap
75	35	87.5	382	4	US-10-425-115-355162	Sequence 355162,
76	35	87.5	391	4	US-10-437-963-204557	Sequence 204557,
77	35	87.5	446	4	US-10-437-963-125297	Sequence 125297,
78	35	87.5	460	4	US-10-425-115-355165	Sequence 355165,
79	35	87.5	481	4	US-10-437-963-173455	Sequence 173455,
80	35	87.5	487	4	US-10-437-963-145352	Sequence 145352,
81	35	87.5	502	4	US-10-437-963-182101	Sequence 182101,
82	35	87.5	507	4	US-10-437-963-128190	Sequence 128190,
83	35	87.5	510	4	US-10-437-963-143565	Sequence 143565,
84	35	87.5	530	4	US-10-437-963-131495	Sequence 131495,
85	35	87.5	567	4	US-10-437-963-184291	Sequence 184291,
86	35	87.5	587	4	US-10-425-114-68483	Sequence 68483, A
87	35	87.5	788	4	US-10-437-963-164365	Sequence 164365,
88	35	87.5	896	4	US-10-174-677-36	Sequence 36, Appl
89	35	87.5	896	4	US-10-367-978-14	Sequence 14, Appl
90	35	87.5	976	4	US-10-437-963-193272	Sequence 193272,
91	35	87.5	1037	5	US-10-450-763-52725	Sequence 52725, A
92	35	87.5	1040	4	US-10-211-462-157	Sequence 157, App
93	35	87.5	1079	3	US-09-820-843A-20	Sequence 20, Appl
94	35	87.5	1093	3	US-09-894-159-71	Sequence 71, Appl
95	34	85.0	13	4	US-10-285-394-159	Sequence 159, App
96	34	85.0	29	4	US-10-424-599-160014	Sequence 160014,
97	34	85.0	58	4	US-10-424-599-200171	Sequence 200171,
98	34	85.0	59	4	US-10-425-115-274584	Sequence 274584,
99	34	85.0	64	4	US-10-437-963-182674	Sequence 182674,
100	34	85.0	69	4	US-10-425-115-332724	Sequence 332724,

ALIGNMENTS

RESULT 1

```

US-10-027-015A-2
; Sequence 2, Application US/10027015A
; Publication No. US2005014312A1
; GENERAL INFORMATION:
; APPLICANT: AKELLA, RAMA
; APPLICANT: RANIERI, JOHN P.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PROMOTING MYOCARDIAL AND PERIPHERAL
; TITLE OF INVENTION: ANGIOGENESIS
; FILE REFERENCE: 2103.000500
; CURRENT APPLICATION NUMBER: US/10/027,015A
; CURRENT FILING DATE: 2001-12-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: SYNTHETIC PEPTIDE
; US-10-027-015A-2

```

```

; Sequence 103936, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 103936
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_101318C.1.pep
US-10-437-963-103936

Query Match          97.5%; Score 39; DB 4; Length 60;
Best Local Similarity 85.7%; Pred. No. 66;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGIGDGG 7
    ||:||||
Db 39 GGVGDGG 45

RESULT 4
US-10-732-923-3741
; Sequence 3741, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 3741
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(88)
; OTHER INFORMATION: unsure at all xaa locations
US-10-732-923-3741

Query Match          97.5%; Score 39; DB 5; Length 88;
Best Local Similarity 85.7%; Pred. No. 94;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGIGDGG 7
    ||:||||
Db 82 GGVGDGG 88

RESULT 5
US-10-437-963-170096
; Sequence 170096, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.

```


; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 170096
; LENGTH: 156
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_68453C.1.1.pep
US-10-437-963-170096

Query Match 97.5%; Score 39; DB 4; Length 156;
Best Local Similarity 85.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
DB 68 GGVGDGG 74

RESULT 6

; Sequence 44299, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 44299
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005919.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.91
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.86
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.87
; OTHER INFORMATION: SWISSPROT HIT: Q99456, EVALUE 2.00e-11
; OTHER INFORMATION: EST_HUMAN HIT: AU118122.1, EVALUE 1.00e-08
US-09-864-761-44299

Query Match 97.5%; Score 39; DB 3; Length 157;
Best Local Similarity 85.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
DB 111 GGVGDGG 117

RESULT 7

US-10-437-963-149979
; Sequence 149979, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 149979
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(184)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_50259C.1.pep
US-10-437-963-149979

Query Match 97.5%; Score 39; DB 4; Length 184;
Best Local Similarity 85.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
DB 137 GGVGDGG 143

RESULT 8

US-09-925-298-517
; Sequence 517, Application US/09925298
; Publication No. US20020039764A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103
; CURRENT APPLICATION NUMBER: US/09/925,298
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 517
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (118)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (161)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (204)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-298-517

Query Match 97.5%; Score 39; DB 3; Length 250;
Best Local Similarity 85.7%; Pred. No. 2.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
||:||||
Db 171 GGVGDGG 177

RESULT 9

US-10-102-806-517
; Sequence 517, Application US/10102806
; Publication No. US20030054421A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103P1C1
; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 517
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (118)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (161)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (204)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-102-806-517

Query Match 97.5%; Score 39; DB 3; Length 250;
Best Local Similarity 85.7%; Pred. No. 2.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
||:||||
Db 171 GGVGDGG 177

US-10-102-806-517

Query Match 97.5%; Score 39; DB 4; Length 250;
Best Local Similarity 85.7%; Pred. No. 2.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
||:||||
Db 171 GGVGDGG 177

RESULT 10

US-10-437-963-135984
; Sequence 135984, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 135984
; LENGTH: 333
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_37606C.1.pep
US-10-437-963-135984

Query Match 97.5%; Score 39; DB 4; Length 333;
Best Local Similarity 85.7%; Pred. No. 3.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
||:||||
Db 6 GGVGDGG 12

RESULT 11

US-10-425-115-277369
; Sequence 277369, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 277369
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_184539C.1.pep
US-10-425-115-277369

Query Match 97.5%; Score 39; DB 4; Length 377;
Best Local Similarity 85.7%; Pred. No. 3.6e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
 ||:||||
 Db 149 GGVGDGG 155

RESULT 12

US-10-450-763-59000
 ; Sequence 59000, Application US/10450763
 ; Publication No. US20050196754A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hyseq, Inc
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
 ; FILE REFERENCE: 790CIP3/US
 ; CURRENT APPLICATION NUMBER: US/10/450,763
 ; CURRENT FILING DATE: 2003-06-11
 ; PRIOR APPLICATION NUMBER: PCT/US01/08631
 ; PRIOR FILING DATE: 2001-03-30
 ; PRIOR APPLICATION NUMBER: 09/540,217
 ; PRIOR FILING DATE: 2000-03-31
 ; PRIOR APPLICATION NUMBER: 09/649,167
 ; PRIOR FILING DATE: 2000-08-23
 ; NUMBER OF SEQ ID NOS: 60736
 ; SOFTWARE: Custom
 ; SEQ ID NO 59000
 ; LENGTH: 399
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; NAME/KEY: DOMAIN
 ; LOCATION: (38)..(50)
 ; OTHER INFORMATION: Channel forming colicins proteins domain identified by
 ; OTHER INFORMATION: eMATRIX, accession number BL00276A, p-value=2.227e-11, raw score
 ; OTHER INFORMATION: 8.87
 US-10-450-763-59000

Query Match 97.5%; Score 39; DB 5; Length 399;
 Best Local Similarity 85.7%; Pred. No. 3.8e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
 ||:||||
 Db 28 GGVGDGG 34

RESULT 13

US-10-437-963-183531
 ; Sequence 183531, Application US/10437963
 ; Publication No. US20040123343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Andrey A.
 ; APPLICANT: Barbazuk, Brad
 ; APPLICANT: Li, Ping
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53221)B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ ID NOS: 204966
 ; SEQ ID NO 183531
 ; LENGTH: 897
 ; TYPE: PRT
 ; ORGANISM: Oryza sativa
 ; NAME/KEY: unsure
 ; LOCATION: (1)..(897)
 ; OTHER INFORMATION: unsure at all Xaa locations
 ; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_80613C.1.pap
 US-10-437-963-183531

Query Match 97.5%; Score 39; DB 4; Length 897;
 Best Local Similarity 85.7%; Pred. No. 8.1e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
 ||:||||
 Db 881 GGVGDGG 887

RESULT 14

US-10-029-386-29024
 ; Sequence 29024, Application US/10029386
 ; Publication No. US20030194704A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharon G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
 ; FILE REFERENCE: AEOMICA-X-2
 ; CURRENT APPLICATION NUMBER: US/10/029,386
 ; CURRENT FILING DATE: 2001-12-20
 ; NUMBER OF SEQ ID NOS: 34288
 ; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 29024
 ; LENGTH: 37
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO AC027045.2
 ; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.61
 ; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.48
 ; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.5
 ; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.59
 ; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
 US-10-029-386-29024

Query Match 95.0%; Score 38; DB 4; Length 37;
 Best Local Similarity 85.7%; Pred. No. 60;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
 ||:||||
 Db 31 GGVGDGG 37

RESULT 15

US-10-437-963-180674
 ; Sequence 180674, Application US/10437963
 ; Publication No. US20040123343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Andrey A.
 ; APPLICANT: Barbazuk, Brad
 ; APPLICANT: Li, Ping
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53221)B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ ID NOS: 204966
 ; SEQ ID NO 180674
 ; LENGTH: 120
 ; TYPE: PRT
 ; ORGANISM: Oryza sativa
 ; NAME/KEY: unsure
 ; LOCATION: (1)..(120)
 ; OTHER INFORMATION: unsure at all Xaa locations
 ; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_78022C.1.pep
US-10-437-963-180674

Query Match 95.0%; Score 38; DB 4; Length 120;
Best Local Similarity 85.7%; Pred. No. 1.8e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
Db 85 GGLGDGG 91

RESULT 16

US-10-425-115-201529
; Sequence 201529, Application US/10425115
; Publication No. US2004014272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 201529
; LENGTH: 122
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_115378C.1.pep
US-10-425-115-201529

Query Match 95.0%; Score 38; DB 4; Length 122;
Best Local Similarity 85.7%; Pred. No. 1.8e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
Db 108 GGLGDGG 114

RESULT 17

US-10-437-963-154584
; Sequence 154584, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 154584
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_5442C.1.pep
US-10-437-963-154584

Query Match 95.0%; Score 38; DB 4; Length 170;
Best Local Similarity 85.7%; Pred. No. 2.5e+02;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
Db 7 GGLGDGG 13

RESULT 18

US-10-437-963-141956
; Sequence 141956, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 141956
; LENGTH: 181
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(181)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_43009C.1.pep
US-10-437-963-141956

Query Match 95.0%; Score 38; DB 4; Length 181;
Best Local Similarity 85.7%; Pred. No. 2.6e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
Db 81 GGLGDGG 87

RESULT 19

US-10-437-963-111588
; Sequence 111588, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 111588
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_15553C.1.pep
US-10-437-963-111588

```
Query Match          95.0%; Score 38; DB 4; Length 214;
Best Local Similarity 85.7%; Pred. No. 3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
DB 98 GGLGDGG 104

RESULT 20
US-10-437-963-157722
; Sequence 157722, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 157722
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_57265C.1.pep
US-10-437-963-157722

Query Match          95.0%; Score 38; DB 4; Length 300;
Best Local Similarity 85.7%; Pred. No. 4.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
DB 188 GGLGDGG 194

RESULT 21
US-11-097-143-9876
; Sequence 9876, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al,
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
```

```
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9876
; LENGTH: 427
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-9876
```

```
Query Match          95.0%; Score 38; DB 6; Length 427;
Best Local Similarity 85.7%; Pred. No. 5.8e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GGIGDGG 7
DB 296 GGLGDGG 302
```

```
RESULT 22
US-10-437-963-173239
; Sequence 173239, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 173239
; LENGTH: 1152
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_71297C.1.pep
US-10-437-963-173239
```

```
Query Match          95.0%; Score 38; DB 4; Length 1152;
Best Local Similarity 85.7%; Pred. No. 1.4e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GGIGDGG 7
DB 313 GGLGDGG 319
```

```
RESULT 23
US-10-437-963-121618
; Sequence 121618, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
```

; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 121618
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_24625C.1.pep
US-10-437-963-121618

Query Match 92.5%; Score 37; DB 4; Length 146;
Best Local Similarity 85.7%; Pred. No. 3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
|||:||||
Db 137 GGMGDGG 143

RESULT 24

US-10-425-115-229793
; Sequence 229793, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 229793
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_141166C.1.pep
US-10-425-115-229793

Query Match 90.0%; Score 36; DB 4; Length 93;
Best Local Similarity 85.7%; Pred. No. 2.8e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
|||:||||
Db 9 GGFGDGG 15

RESULT 25

US-10-425-115-205711
; Sequence 205711, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 205711
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_119193C.1.pep
US-10-425-115-205711

Query Match 90.0%; Score 36; DB 4; Length 117;
Best Local Similarity 85.7%; Pred. No. 3.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
|||:||||
Db 101 GGFGDGG 107

RESULT 26

US-10-437-963-196524
; Sequence 196524, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 196524
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(170)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_92368C.1.pep
US-10-437-963-196524

Query Match 90.0%; Score 36; DB 4; Length 170;
Best Local Similarity 85.7%; Pred. No. 5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
|||:||||
Db 9 GGIGEGG 15

RESULT 27

US-10-437-963-130467
; Sequence 130467, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 130467
; LENGTH: 187
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:

; NAME/KEY: unsure
; LOCATION: (1)..(187)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_32626C.1.pep
US-10-437-963-130467

Query Match 90.0%; Score 36; DB 4; Length 187;
Best Local Similarity 85.7%; Pred. No. 5.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGIGDGG 7
|||:||||
Db 2 GGIGEGG 8

RESULT 28

US-10-437-963-159222
; Sequence 159222, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 159222
; LENGTH: 290
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(290)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_5861C.1.pep
US-10-437-963-159222

Query Match 90.0%; Score 36; DB 4; Length 290;
Best Local Similarity 85.7%; Pred. No. 8.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGIGDGG 7
|||:||||
Db 215 GGFGDGG 221

RESULT 29

US-10-425-115-211110
; Sequence 211110, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 211110
; LENGTH: 292

; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(292)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_124129C.1.pep
US-10-425-115-211110

Query Match 90.0%; Score 36; DB 4; Length 292;
Best Local Similarity 85.7%; Pred. No. 8.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGIGDGG 7
|||:||||
Db 27 GGFGDGG 33

RESULT 30

US-10-437-963-149137
; Sequence 149137, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 149137
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_494C.1.pep
US-10-437-963-149137

Query Match 90.0%; Score 36; DB 4; Length 383;
Best Local Similarity 85.7%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGIGDGG 7
|||:||||
Db 112 GGFGDGG 118

RESULT 31

US-10-437-963-172290
; Sequence 172290, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 172290
; LENGTH: 405
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_70440C.1.pap
US-10-437-963-172290

Query Match 90.0%; Score 36; DB 4; Length 405;
Best Local Similarity 85.7%; Pred. No. 1.1e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
|||:|
Db 66 GGIGGG 72

RESULT 32
US-10-156-761-10244
; Sequence 10244, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 10244
; LENGTH: 456
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-10244

Query Match 90.0%; Score 36; DB 4; Length 456;
Best Local Similarity 85.7%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
|||:|
Db 434 GGFGDGG 440

RESULT 33
US-10-437-963-159271
; Sequence 159271, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 159271
; LENGTH: 813
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_58664C.1.pap
US-10-437-963-159271

Query Match 90.0%; Score 36; DB 4; Length 813;
Best Local Similarity 85.7%; Pred. No. 2.1e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
|||:|
Db 53 GGFGDGG 59

RESULT 34
US-10-156-761-10231
; Sequence 10231, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 10231
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-10231

Query Match 90.0%; Score 36; DB 4; Length 1278;
Best Local Similarity 85.7%; Pred. No. 3.2e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
|||:|
Db 848 GGFGDGG 854

RESULT 35
US-10-203-295-7
; Sequence 7, Application US/10203295
; Publication No. US20040115762A1
; GENERAL INFORMATION:
; APPLICANT: Zotchev, Sergey Borisovich
; APPLICANT: Sekurova, Olga Nikalayivna
; APPLICANT: Fjaervik, Epsen
; APPLICANT: Brautaset, Trygve
; APPLICANT: Strom, Arne Reidar
; APPLICANT: Valla, Svein
; APPLICANT: Ellingsen, Trond Erling
; APPLICANT: Sletta, Havard
; APPLICANT: Gulliksen, Ole-Martin
; TITLE OF INVENTION: Novel genes encoding a nystatin polyketide synthase and their
; FILE OF INVENTION: manipulation and utility
; FILE REFERENCE: 1181-265
; CURRENT APPLICATION NUMBER: US/10/203,295
; CURRENT FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: PCT/GB 01/00509


```

; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: GB 0002840.7
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: GB 0008786.6
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: GB 0009387.2
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 11088
; TYPE: PRT
; ORGANISM: Streptomyces noursei ATCC 11455
US-10-203-295-7

Query Match
Best Local Similarity 90.0%; Score 36; DB 4; Length 11088;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGIGDGG 7
Db 8417 GGFGDGG 8423

RESULT 36
US-10-732-923-20557
; Sequence 20557, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 20557
; LENGTH: 11096
; TYPE: PRT
; ORGANISM: Streptomyces noursei
US-10-732-923-20557

Query Match
Best Local Similarity 90.0%; Score 36; DB 5; Length 11096;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGIGDGG 7
Db 8425 GGFGDGG 8431

RESULT 37
US-09-864-761-36444
; Sequence 36444, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6

```

```

; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 36444
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC010099.3
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2
US-09-864-761-36444

Query Match
Best Local Similarity 87.5%; Score 35; DB 3; Length 61;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGIGDGG 7
Db 32 GGFGDGG 38

RESULT 38
US-09-864-761-42760
; Sequence 42760, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312

```

; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 42760
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC010099.2
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.95
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.97
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.92
US-09-864-761-42760

Query Match 87.5%; Score 35; DB 3; Length 61;
Best Local Similarity 85.7%; Pred. No. 2.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
|||
Db 32 GGYGDGG 38

RESULT 39
US-10-424-599-192048
; Sequence 192048, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 192048
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_15440C.1.pap
US-10-424-599-192048

Query Match 87.5%; Score 35; DB 4; Length 70;
Best Local Similarity 85.7%; Pred. No. 3.1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
|||
Db 35 GGYGDGG 41

RESULT 40
US-10-437-963-109585
; Sequence 109585, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 109585
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_13728C.1.pap
US-10-437-963-109585

Query Match 87.5%; Score 35; DB 4; Length 75;
Best Local Similarity 85.7%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
|||
Db 58 GGYGDGG 64

RESULT 41
US-10-437-963-142504
; Sequence 142504, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 142504
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_43504C.1.pep
US-10-437-963-142504

Query Match 87.5%; Score 35; DB 4; Length 84;
Best Local Similarity 85.7%; Pred. No. 3.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
|||
Db 60 GGTGDGG 66

RESULT 42
US-10-424-599-184457
; Sequence 184457, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 184457
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(86)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_137580C.1.pep
US-10-424-599-184457

Query Match 87.5%; Score 35; DB 4; Length 86;
Best Local Similarity 85.7%; Pred. No. 3.8e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
|||
Db 19 GGAGDGG 25

RESULT 43
US-10-425-115-193758
; Sequence 193758, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 193758
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:

; NAME/KEY: unsure
; LOCATION: (1)..(92)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_10828C.1.pep
US-10-425-115-193758

Query Match 87.5%; Score 35; DB 4; Length 92;
Best Local Similarity 85.7%; Pred. No. 4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
|||
Db 77 GGAGDGG 83

RESULT 44
US-10-437-963-153974
; Sequence 153974, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 153974
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_53878C.1.pep
US-10-437-963-153974

Query Match 87.5%; Score 35; DB 4; Length 96;
Best Local Similarity 85.7%; Pred. No. 4.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
|||
Db 70 GGAGDGG 76

RESULT 45
US-10-437-963-129822
; Sequence 129822, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 129822
; LENGTH: 97

```
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(97)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_32041C.1.pep
US-10-437-963-129822

Query Match      87.5%; Score 35; DB 4; Length 97;
Best Local Similarity 85.7%; Pred. No. 4.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
   |||||
Db 31 GGAGDGG 37

RESULT 46
US-10-767-701-42596
; Sequence 42596, Application US/10767701
; Publication No. US20040172694A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 42596
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(100)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C74510_1.pep
US-10-767-701-42596

Query Match      87.5%; Score 35; DB 4; Length 100;
Best Local Similarity 85.7%; Pred. No. 4.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
   |||||
Db 56 GGYGDGG 62

RESULT 47
US-10-425-115-189769
; Sequence 189769, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 189769
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Zea mays
```

```
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(108)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_104653C.1.pep
US-10-425-115-189769

Query Match      87.5%; Score 35; DB 4; Length 108;
Best Local Similarity 85.7%; Pred. No. 4.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
   |||||
Db 18 GGAGDGG 24

RESULT 48
US-10-424-599-200912
; Sequence 200912, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 200912
; LENGTH: 119
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_23449C.1.pep
US-10-424-599-200912

Query Match      87.5%; Score 35; DB 4; Length 119;
Best Local Similarity 71.4%; Pred. No. 5.1e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
   ||:|
Db 88 GGVGEGG 94

RESULT 49
US-10-437-963-140184
; Sequence 140184, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 140184
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
```

; NAME/KEY: unsure
; LOCATION: (1)..(128)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_41406C.1.pap
US-10-437-963-140184

Query Match 87.5%; Score 35; DB 4; Length 128;
Best Local Similarity 85.7%; Pred. No. 5.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGIGDGG 7
Db 35 GGAGDGG 41

RESULT 50
US-10-425-115-188673
; Sequence 188673, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 188673
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(132)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_103659C.1.pap
US-10-425-115-188673

Query Match 87.5%; Score 35; DB 4; Length 132;
Best Local Similarity 85.7%; Pred. No. 5.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGIGDGG 7
Db 122 GGAGDGG 128

Search completed: February 27, 2006, 10:04:33
Job time : 167 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (C) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 27, 2006, 10:01:55 ; Search time 18 Seconds
(without alignments)

5.789 Million cell updates/sec

Title: US-10-027-015B-2

Perfect score: 40

Sequence: 1 GGIGDGG 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 117670 seqs, 14887254 residues

Total number of hits satisfying chosen parameters: 117670

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Published Applications AA New:*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	97.5	544	US-10-821-234-889	Sequence 889, App
2	35	87.5	228	US-10-714-887-90	Sequence 90, Appl
3	35	87.5	233	US-10-714-887-100	Sequence 100, App
4	35	87.5	1079	US-11-052-554A-145	Sequence 145, App
5	34	85.0	669	US-11-024-959-520	Sequence 520, App
6	34	85.0	731	US-11-052-554A-155	Sequence 155, App
7	34	85.0	806	US-11-108-172-1117	Sequence 1117, App
8	34	85.0	806	US-11-108-172-1118	Sequence 1118, App
9	34	85.0	806	US-11-079-900-1	Sequence 1, Appli
10	33	82.5	131	US-11-072-512-3039	Sequence 3039, Ap
11	33	82.5	152	US-10-467-657-7086	Sequence 7086, Ap
12	33	82.5	394	US-11-057-012-92	Sequence 92, Appl
13	33	82.5	398	US-11-057-012-62	Sequence 62, Appl
14	33	82.5	415	US-11-057-012-64	Sequence 64, Appl
15	33	82.5	426	US-11-230-180-9	Sequence 9, Appli
16	33	82.5	621	US-10-131-826A-40	Sequence 40, Appl
17	33	82.5	648	US-11-024-959-457	Sequence 457, App
18	33	82.5	802	US-11-037-243-113	Sequence 113, App
19	33	82.5	1660	US-11-052-554A-137	Sequence 137, App
20	33	82.5	3375	US-11-044-111-23	Sequence 23, Appl
21	32	80.0	91	US-10-467-657-3590	Sequence 3590, Ap
22	32	80.0	91	US-10-467-657-6440	Sequence 6440, Ap
23	32	80.0	290	US-10-714-887-110	Sequence 110, App
24	32	80.0	434	US-11-052-554A-167	Sequence 167, App
25	32	80.0	541	US-11-059-814-20	Sequence 20, Appl

26	32	80.0	558	6	US-10-504-364-3	Sequence 3, Appli
27	32	80.0	558	6	US-10-504-364-4	Sequence 4, Appli
28	32	80.0	618	7	US-11-052-554A-150	Sequence 150, App
29	32	80.0	710	7	US-11-045-802-22	Sequence 22, Appl
30	32	80.0	710	7	US-11-045-802-23	Sequence 23, Appl
31	32	80.0	767	7	US-11-052-554A-154	Sequence 154, App
32	32	80.0	891	7	US-11-205-109-18	Sequence 18, Appl
33	32	80.0	923	7	US-11-052-554A-147	Sequence 147, App
34	31	77.5	212	6	US-10-523-503-44	Sequence 44, Appl
35	31	77.5	265	6	US-10-714-887-232	Sequence 232, App
36	31	77.5	270	6	US-10-467-657-5426	Sequence 5426, Ap
37	31	77.5	430	7	US-11-124-367A-460	Sequence 460, App
38	31	77.5	530	7	US-11-110-082-32	Sequence 32, Appl
39	31	77.5	547	7	US-11-072-175-243	Sequence 243, App
40	31	77.5	602	7	US-11-072-512-3324	Sequence 3324, Ap
41	31	77.5	606	7	US-11-052-554A-163	Sequence 163, App
42	31	77.5	659	6	US-10-661-966-14	Sequence 14, Appl
43	31	77.5	1381	7	US-11-052-554A-138	Sequence 138, App
44	31	77.5	1901	7	US-11-052-554A-135	Sequence 135, App
45	30	75.0	16	6	US-10-939-890-492	Sequence 492, App
46	30	75.0	189	6	US-10-467-657-922	Sequence 922, App
47	30	75.0	200	6	US-10-467-657-2106	Sequence 2106, Ap
48	30	75.0	229	6	US-10-883-512-88	Sequence 88, Appl
49	30	75.0	275	7	US-11-232-455A-40	Sequence 40, Appl
50	30	75.0	324	6	US-10-714-887-300	Sequence 300, App
51	30	75.0	353	6	US-10-467-657-3964	Sequence 3964, Ap
52	30	75.0	384	7	US-11-098-686-10625	Sequence 10625, A
53	30	75.0	644	7	US-11-072-512-2945	Sequence 2945, Ap
54	30	75.0	749	7	US-11-052-554A-148	Sequence 148, App
55	30	75.0	853	7	US-11-052-554A-149	Sequence 149, App
56	30	75.0	1618	6	US-10-984-645-2	Sequence 2, Appli
57	29.5	73.8	402	6	US-10-467-657-6870	Sequence 6870, Ap
58	29.5	73.8	443	6	US-10-467-657-3768	Sequence 3768, Ap
59	29.5	73.8	443	6	US-10-467-657-6920	Sequence 6920, Ap
60	29.5	73.8	837	7	US-11-052-554A-159	Sequence 159, App
61	29	72.5	43	6	US-10-467-657-5420	Sequence 5420, Ap
62	29	72.5	51	7	US-11-088-785-13	Sequence 13, Appl
63	29	72.5	56	7	US-11-088-785-9	Sequence 9, Appli
64	29	72.5	56	7	US-11-088-785-10	Sequence 10, Appl
65	29	72.5	56	7	US-11-088-785-12	Sequence 12, Appl
66	29	72.5	61	7	US-11-088-785-11	Sequence 11, Appl
67	29	72.5	95	6	US-10-467-657-8785	Sequence 8785, Ap
68	29	72.5	101	7	US-11-110-424-1	Sequence 1, Appli
69	29	72.5	101	7	US-11-110-424-2	Sequence 2, Appli
70	29	72.5	101	7	US-11-110-424-3	Sequence 3, Appli
71	29	72.5	107	6	US-10-467-657-4460	Sequence 4460, Ap
72	29	72.5	116	7	US-11-087-227-78	Sequence 78, Appl
73	29	72.5	361	7	US-11-052-554A-169	Sequence 169, App
74	29	72.5	392	6	US-10-793-626-2794	Sequence 2794, Ap
75	29	72.5	392	6	US-10-793-626-2874	Sequence 2874, Ap
76	29	72.5	410	6	US-10-467-657-2742	Sequence 2742, Ap
77	29	72.5	422	6	US-10-917-905-3	Sequence 3, Appli
78	29	72.5	433	7	US-11-167-856-26	Sequence 26, Appl
79	29	72.5	567	7	US-11-072-512-3772	Sequence 3772, Ap
80	29	72.5	617	7	US-11-072-512-2915	Sequence 2915, Ap
81	29	72.5	639	7	US-11-052-554A-165	Sequence 165, App
82	29	72.5	778	7	US-11-052-554A-144	Sequence 144, App
83	29	72.5	910	7	US-11-088-785-5	Sequence 5, Appli
84	29	72.5	1052	6	US-10-467-657-3992	Sequence 3992, Ap
85	29	72.5	1329	7	US-11-052-554A-136	Sequence 136, App
86	29	72.5	2011	7	US-11-080-991-56	Sequence 56, Appl
87	29	72.5	3157	7	US-11-052-554A-152	Sequence 152, App
88	28.5	71.2	615	7	US-11-052-554A-142	Sequence 142, App
89	28.5	71.2	694	7	US-11-052-554A-158	Sequence 158, App
90	28.5	71.2	741	7	US-11-052-554A-161	Sequence 161, App
91	28.5	71.2	1538	7	US-11-052-554A-146	Sequence 146, App
92	28	70.0	11	6	US-10-957-351-346	Sequence 346, App
93	28	70.0	12	7	US-11-065-943-39	Sequence 39, Appl
94	28	70.0	15	6	US-10-861-105-1	Sequence 1, Appli
95	28	70.0	15	6	US-11-056-182-17	Sequence 17, Appl
96	28	70.0	17	6	US-10-957-351-339	Sequence 339, App
97	28	70.0	17	6	US-10-957-351-340	Sequence 340, App
98	28	70.0	18	7	US-11-115-922-243	Sequence 243, App

99 28 70.0 19 6 US-10-948-053-15 Sequence 15, Appl
100 28 70.0 25 6 US-10-957-351-345 Sequence 345, App

ALIGNMENTS

RESULT 1
US-10-821-234-889
; Sequence 889, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: PC_SEQ_Genes Version 1.0
; SEQ ID NO 889
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-889

Query Match 97.5%; Score 39; DB 6; Length 544;
Best Local Similarity 85.7%; Pred. No. 27;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
||:||||
Db 147 GGVGDGG 153

RESULT 2
US-10-714-887-90
; Sequence 90, Application US/10714887
; Publication No. US20060015972A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: HEARD, Jacqueline
; APPLICANT: RIECHMANN, Jose Luis
; APPLICANT: CREELMAN, Robert
; APPLICANT: RATCLIFFE, Oliver
; APPLICANT: CANALES, Roger
; APPLICANT: REPETTI, Peter
; APPLICANT: KUMIMOTO, Roderick W
; APPLICANT: GUTTERSON, Neal
; APPLICANT: REUBER, T. Lynne
; APPLICANT: PINEDA, Omaira
; APPLICANT: SHERMAN, Bradley K
; TITLE OF INVENTION: PLANT TRANSCRIPTIONAL REGULATORS OF DROUGHT STRESS
; FILE REFERENCE: MB10058-CIP
; CURRENT APPLICATION NUMBER: US/10/714,887
; CURRENT FILING DATE: 2003-11-13
; PRIOR APPLICATION NUMBER: 10/412,699
; PRIOR FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: 09/506,720
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/135,134
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: 09/394,519
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: 09/533,392
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,029
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/532,591
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,030
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/125,814
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: 09/713,994
; PRIOR FILING DATE: 2000-11-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 430
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 100
; LENGTH: 233

; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,030
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/125,814
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: 09/713,994
; PRIOR FILING DATE: 2000-11-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 430
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 90
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Oryza sativa (japonica cultivar-group)
; FEATURE:
; OTHER INFORMATION: G3644 polypeptide Orthologous to G47
US-10-714-887-90

Query Match 87.5%; Score 35; DB 6; Length 228;
Best Local Similarity 85.7%; Pred. No. 49;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
||:||||
Db 104 GGVGDGG 110

RESULT 3
US-10-714-887-100
; Sequence 100, Application US/10714887
; Publication No. US20060015972A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: HEARD, Jacqueline
; APPLICANT: RIECHMANN, Jose Luis
; APPLICANT: CREELMAN, Robert
; APPLICANT: RATCLIFFE, Oliver
; APPLICANT: CANALES, Roger
; APPLICANT: REPETTI, Peter
; APPLICANT: KUMIMOTO, Roderick W
; APPLICANT: GUTTERSON, Neal
; APPLICANT: REUBER, T. Lynne
; APPLICANT: PINEDA, Omaira
; APPLICANT: SHERMAN, Bradley K
; TITLE OF INVENTION: PLANT TRANSCRIPTIONAL REGULATORS OF DROUGHT STRESS
; FILE REFERENCE: MB10058-CIP
; CURRENT APPLICATION NUMBER: US/10/714,887
; CURRENT FILING DATE: 2003-11-13
; PRIOR APPLICATION NUMBER: 10/412,699
; PRIOR FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: 09/506,720
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/135,134
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: 09/394,519
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: 09/533,392
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,029
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/532,591
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,030
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/125,814
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: 09/713,994
; PRIOR FILING DATE: 2000-11-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 430
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 100
; LENGTH: 233


```
; TYPE: PRT
; ORGANISM: Oryza sativa (japonica cultivar-group)
; FEATURE:
; OTHER INFORMATION: G3651 polypeptide Orthologous to G47
US-10-714-887-100

Query Match      87.5%; Score 35; DB 6; Length 233;
Best Local Similarity 85.7%; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
Db 112 GGAGDGG 118

RESULT 4
US-11-052-554A-145
; Sequence 145, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 145
; LENGTH: 1079
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis H37Rv
US-11-052-554A-145

Query Match      87.5%; Score 35; DB 7; Length 1079;
Best Local Similarity 85.7%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
Db 1048 GGAGDGG 1054

RESULT 5
US-11-024-959-520
; Sequence 520, Application US/11024959
; Publication No. US20060010516A1
; GENERAL INFORMATION:
; APPLICANT: FORSTER, RICHARD L.
; APPLICANT: CONNETT, MARIE B.
; APPLICANT: EMERSON, SARAH JANE
; APPLICANT: GRIGOR, MURRAY ROBERT
; APPLICANT: HIGGINS, COLLEEN M.
; APPLICANT: LUND, STEVEN TROY
; APPLICANT: MAGUSIN, ANDREAS
; APPLICANT: KODRZYCKI, BOB
; TITLE OF INVENTION: CELL CYCLE GENES AND RELATED METHODS
; FILE REFERENCE: 044463-0360
; CURRENT APPLICATION NUMBER: US/11/024,959
; CURRENT FILING DATE: 2004-12-30
; PRIOR APPLICATION NUMBER: 60/533,036
; PRIOR FILING DATE: 2003-12-30
; NUMBER OF SEQ ID NOS: 782
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 520
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Eucalyptus sp.
US-11-024-959-520
```

```
Query Match      85.0%; Score 34; DB 7; Length 669;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 6
Db 370 GGIGDGG 375

RESULT 6
US-11-052-554A-155
; Sequence 155, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 155
; LENGTH: 731
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis H37Rv
US-11-052-554A-155

Query Match      85.0%; Score 34; DB 7; Length 731;
Best Local Similarity 71.4%; Pred. No. 2.1e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
Db 438 GGVGNGG 444

RESULT 7
US-11-108-172-1117
; Sequence 1117, Application US/11108172
; Publication No. US20050260177A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole L.
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick Thomas S.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.471C15
; CURRENT APPLICATION NUMBER: US/11/108,172
; CURRENT FILING DATE: 2005-04-15
; PRIOR APPLICATION NUMBER: US 10/025,380
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: US 09/922,217
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 09/833,263
```

```
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 09/649,811
; PRIOR FILING DATE: 2000-08-28
; PRIOR APPLICATION NUMBER: US 09/609,448
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: US 09/575,251
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/519,444
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 09/504,629
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: US 09/480,321
; PRIOR FILING DATE: 2000-01-10
; PRIOR APPLICATION NUMBER: US 09/476,296
; PRIOR FILING DATE: 1999-12-30
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1130
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1117
; LENGTH: 806
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-108-172-1117
```

```
Query Match      85.0%; Score 34; DB 7; Length 806;
Best Local Similarity 85.7%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 GGIGDGG 7
Db      588 GNIGDGG 594
```

RESULT 8

```
US-11-108-172-1118
; Sequence 1118, Application US/11108172
; Publication No. US20050260177A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Smith, Carole L.
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aljun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C15
; CURRENT APPLICATION NUMBER: US/11/108,172
; CURRENT FILING DATE: 2005-04-15
; PRIOR APPLICATION NUMBER: US 10/025,380
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: US 09/922,217
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 09/833,263
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 09/649,811
; PRIOR FILING DATE: 2000-08-28
; PRIOR APPLICATION NUMBER: US 09/609,448
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: US 09/575,251
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/519,444
; PRIOR FILING DATE: 2000-03-06
```

```
; PRIOR APPLICATION NUMBER: US 09/504,629
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: US 09/480,321
; PRIOR FILING DATE: 2000-01-10
; PRIOR APPLICATION NUMBER: US 09/476,296
; PRIOR FILING DATE: 1999-12-30
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1130
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1118
; LENGTH: 806
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-108-172-1118
```

```
Query Match      85.0%; Score 34; DB 7; Length 806;
Best Local Similarity 85.7%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 GGIGDGG 7
Db      588 GNIGDGG 594
```

RESULT 9

```
US-11-079-900-1
; Sequence 1, Application US/11079900
; Publication No. US20060014168A1
; GENERAL INFORMATION:
; APPLICANT: KIMONIS, VIRGINIA E.
; APPLICANT: WATTS, GILES D. J.
; TITLE OF INVENTION: METHOD OF DIAGNOSIS OF INCLUSION BODY MYOPATHY-PAGET
; TITLE OF INVENTION: BONE DISEASE-FRONTOTEMPORAL DEMENTIA SYNDROME
; FILE REFERENCE: 701039-55081-US
; CURRENT APPLICATION NUMBER: US/11/079,900
; CURRENT FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: 60/552,562
; PRIOR FILING DATE: 2004-03-12
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 1
; LENGTH: 806
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-079-900-1
```

```
Query Match      85.0%; Score 34; DB 7; Length 806;
Best Local Similarity 85.7%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 GGIGDGG 7
Db      588 GNIGDGG 594
```

RESULT 10

```
US-11-072-512-3039
; Sequence 3039, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
```

APPLICANT: SEKI, NAOHIKO
APPLICANT: YOSHIKAWA, TSUTOMU
APPLICANT: OTSUKA, MOTOTYUKI
APPLICANT: NAGAHARI, KENJI
APPLICANT: MASUHO, YASUHIKO
FILE OF INVENTION: Novel full length cDNA
CURRENT APPLICATION NUMBER: US/11/072,512
CURRENT FILING DATE: 2005-03-07
PRIOR APPLICATION NUMBER: US 60/350,978
PRIOR FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: JP 2001-379298
PRIOR FILING DATE: 2001-11-05
NUMBER OF SEQ ID NOS: 4096
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3039
LENGTH: 131
TYPE: PRT
ORGANISM: Homo sapiens
US-11-072-512-3039

Query Match 82.5%; Score 33; DB 7; Length 131;
Best Local Similarity 71.4%; Pred. No. 60;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGIGDGG 7
|||
Db 21 GGVADGG 27

RESULT 11

US-10-467-657-7086
Sequence 7086, Application US/10467657
Publication No. US20050260581A1
GENERAL INFORMATION:
APPLICANT: CHIRON SpA
APPLICANT: FONTANA Maria Rita
APPLICANT: PIZZA Mariagrazia
APPLICANT: MASNANI Vega
APPLICANT: MONACI Elisabetta
TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/467,657
CURRENT FILING DATE: 2003-08-11
PRIOR APPLICATION NUMBER: GB-0103424.8
PRIOR FILING DATE: 2001-02-12
NUMBER OF SEQ ID NOS: 9218
SOFTWARE: SeqWin99, version 1.04
SEQ ID NO 7086
LENGTH: 152
TYPE: PRT
ORGANISM: Neisseria gonorrhoeae
US-10-467-657-7086

Query Match 82.5%; Score 33; DB 6; Length 152;
Best Local Similarity 85.7%; Pred. No. 68;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGIGDGG 7
|||
Db 127 GGRGDGG 133

RESULT 12

US-11-057-012-92
Sequence 92, Application US/11057012
Publication No. US20060026705A1
GENERAL INFORMATION:
APPLICANT: Allen, Stephen M.
APPLICANT: Kinney, Stephen J.
APPLICANT: Liu, Zhan-Bin
APPLICANT: Stecca, Kevin L.
TITLE OF INVENTION: Plant Amino Acid Biosynthetic Enzymes

FILE REFERENCE: BB1116 US CIP - 1
CURRENT APPLICATION NUMBER: US/11/057,012
CURRENT FILING DATE: 2005-02-11
PRIOR APPLICATION NUMBER: 09/931,457
PRIOR FILING DATE: 2001-08-16
PRIOR APPLICATION NUMBER: 09/424,976
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: 60/065,385
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/049,406
PRIOR FILING DATE: 1997-06-12
NUMBER OF SEQ ID NOS: 94
SOFTWARE: Microsoft Office 97
SEQ ID NO 92
LENGTH: 394
TYPE: PRT
ORGANISM: Glycine max
PUBLICATION INFORMATION:
TITLE: New Recombinant DNA Construct
DATABASE ACCESSION NUMBER: USPTAALL / SEQ ID NO:129998
DATABASE ENTRY DATE: 2004-10-07
PATENT DOCUMENT NUMBER: US2004123343
PATENT FILING DATE: 2000-04-19
PUBLICATION DATE: 2004-06-24
RELEVANT RESIDUES: (1) .. (394)
US-11-057-012-92

Query Match 82.5%; Score 33; DB 7; Length 394;
Best Local Similarity 85.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGIGDGG 7
|||||
Db 250 GGIGTGG 256

RESULT 13

US-11-057-012-62
Sequence 62, Application US/11057012
Publication No. US20060026705A1
GENERAL INFORMATION:
APPLICANT: Allen, Stephen M.
APPLICANT: Kinney, Anthony J.
APPLICANT: Liu, Zhan-Bin
APPLICANT: Stecca, Kevin L.
TITLE OF INVENTION: Plant Amino Acid Biosynthetic Enzymes
FILE REFERENCE: BB1116 US CIP - 1
CURRENT APPLICATION NUMBER: US/11/057,012
CURRENT FILING DATE: 2005-02-11
PRIOR APPLICATION NUMBER: 09/931,457
PRIOR FILING DATE: 2001-08-16
PRIOR APPLICATION NUMBER: 09/424,976
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: 60/065,385
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/049,406
PRIOR FILING DATE: 1997-06-12
NUMBER OF SEQ ID NOS: 94
SOFTWARE: Microsoft Office 97
SEQ ID NO 62
LENGTH: 398
TYPE: PRT
ORGANISM: Zea mays
US-11-057-012-62

Query Match 82.5%; Score 33; DB 7; Length 398;
Best Local Similarity 85.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGIGDGG 7
|||||
Db 254 GGIGTGG 260

RESULT 14
 US-11-057-012-64
 ; Sequence 64, Application US/11057012
 ; Publication No. US20060026705A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Stephen M.
 ; APPLICANT: Kinney, Anthony J.
 ; APPLICANT: Liu, Zhan-Bin
 ; APPLICANT: Stecca, Kevin L.
 ; TITLE OF INVENTION: Plant Amino Acid Biosynthetic Enzymes
 ; FILE REFERENCE: BB1116 US CIP - 1
 ; CURRENT APPLICATION NUMBER: US/11/057,012
 ; CURRENT FILING DATE: 2005-02-11
 ; PRIOR APPLICATION NUMBER: 09/931,457
 ; PRIOR FILING DATE: 2001-08-16
 ; PRIOR APPLICATION NUMBER: 09/424,976
 ; PRIOR FILING DATE: 1999-12-02
 ; PRIOR APPLICATION NUMBER: 60/065,385
 ; PRIOR FILING DATE: 1997-11-12
 ; PRIOR APPLICATION NUMBER: 60/049,406
 ; PRIOR FILING DATE: 1997-06-12
 ; NUMBER OF SEQ ID NOS: 94
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 64
 ; LENGTH: 415
 ; TYPE: PRT
 ; ORGANISM: Oryza sativa
 US-11-057-012-64

Query Match 82.5%; Score 33; DB 7; Length 415;
 Best Local Similarity 85.7%; Pred. No. 1.7e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
 |||||
 Db 271 GGIGTGG 277

RESULT 15
 US-11-230-180-9
 ; Sequence 9, Application US/11230180
 ; Publication No. US20060035263A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Woolf, Clifford J.
 ; APPLICANT: Samad, Tarek A.
 ; TITLE OF INVENTION: DRG11-RESPONSIVE (DRAGON) GENE FAMILY
 ; FILE REFERENCE: 00786/419002
 ; CURRENT APPLICATION NUMBER: US/11/230,180
 ; CURRENT FILING DATE: 2005-09-19
 ; PRIOR APPLICATION NUMBER: US/10/419,296
 ; PRIOR FILING DATE: 2003-04-17
 ; PRIOR APPLICATION NUMBER: 60/373,519
 ; PRIOR FILING DATE: 2002-04-18
 ; NUMBER OF SEQ ID NOS: 32
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 9
 ; LENGTH: 426
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 US-11-230-180-9

Query Match 82.5%; Score 33; DB 7; Length 426;
 Best Local Similarity 71.4%; Pred. No. 1.8e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
 |||||
 Db 72 GGVGSGG 78

RESULT 16
 US-10-131-826A-40

; Sequence 40, Application US/10131826A
 ; Publication No. US20050245730A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen
 ; APPLICANT: DeForge, Laura
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3330RIC128
 ; CURRENT APPLICATION NUMBER: US/10/131,826A
 ; CURRENT FILING DATE: 2002-04-24
 ; PRIOR APPLICATION NUMBER: 60/049911
 ; PRIOR FILING DATE: 1997-06-18
 ; PRIOR APPLICATION NUMBER: 60/056974
 ; PRIOR FILING DATE: 1997-08-26
 ; PRIOR APPLICATION NUMBER: 60/059113
 ; PRIOR FILING DATE: 1997-09-17
 ; PRIOR APPLICATION NUMBER: 60/059115
 ; PRIOR FILING DATE: 1997-09-17
 ; PRIOR APPLICATION NUMBER: 60/059117
 ; PRIOR FILING DATE: 1997-09-17
 ; PRIOR APPLICATION NUMBER: 60/059122
 ; PRIOR FILING DATE: 1997-09-17
 ; PRIOR APPLICATION NUMBER: 60/059184
 ; PRIOR FILING DATE: 1997-09-17
 ; PRIOR APPLICATION NUMBER: 60/059263
 ; PRIOR FILING DATE: 1997-09-18
 ; PRIOR APPLICATION NUMBER: 60/059352
 ; PRIOR FILING DATE: 1997-09-19
 ; PRIOR APPLICATION NUMBER: 60/059588
 ; PRIOR FILING DATE: 1997-09-19
 ; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 550
 ; SEQ ID NO 40
 ; LENGTH: 621
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 US-10-131-826A-40

Query Match 82.5%; Score 33; DB 6; Length 621;
 Best Local Similarity 83.3%; Pred. No. 2.5e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 GIGDGG 7
 |||||
 Db 482 GVGDDG 487

RESULT 17
 US-11-024-959-457
 ; Sequence 457, Application US/11024959
 ; Publication No. US20060010516A1
 ; GENERAL INFORMATION:
 ; APPLICANT: FORSTER, RICHARD L.
 ; APPLICANT: CONNETT, MARIE B.
 ; APPLICANT: EMERSON, SARAH JANE
 ; APPLICANT: GRIGOR, MURRAY ROBERT
 ; APPLICANT: HIGGINS, COLLEEN M.
 ; APPLICANT: LUND, STEVEN TROY

; APPLICANT: MAGUSIN, ANDREAS
; APPLICANT: KODZYCKI, BOB
; TITLE OF INVENTION: CELL CYCLE GENES AND RELATED METHODS
; FILE REFERENCE: 044463-0360
; CURRENT APPLICATION NUMBER: US/11/024,959
; CURRENT FILING DATE: 2004-12-30
; PRIOR APPLICATION NUMBER: 60/533,036
; PRIOR FILING DATE: 2003-12-30
; NUMBER OF SEQ ID NOS: 782
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 457
; LENGTH: 648
; TYPE: PRT
; ORGANISM: Pinus radiata
US-11-024-959-457

Query Match 82.5%; Score 33; DB 7; Length 648;
Best Local Similarity 83.3%; Pred. No. 2.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGIGDG 6
Db 348 GGVGDG 353

RESULT 18

US-11-037-243-113
; Sequence 113, Application US/11037243
; Publication No. US20050287546A1
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: WHYTE, DAVID
; APPLICANT: CAENEPEEL, SEAN
; APPLICANT: CHARYDCZAK, GLEN
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1214
; CURRENT APPLICATION NUMBER: US/11/037,243
; CURRENT FILING DATE: 2005-05-26
; PRIOR APPLICATION NUMBER: US/09/888,615
; PRIOR FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/214,047
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 113
; LENGTH: 802
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-037-243-113

Query Match 82.5%; Score 33; DB 7; Length 802;
Best Local Similarity 85.7%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGIGDG 7
Db 11 GGQGDG 17

RESULT 19

US-11-052-554A-137
; Sequence 137, Application US/11052554A
; Publication No. US2005028866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227

; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 137
; LENGTH: 1660
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis H37Rv
US-11-052-554A-137

Query Match 82.5%; Score 33; DB 7; Length 1660;
Best Local Similarity 85.7%; Pred. No. 6.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGIGDG 7
Db 1253 GDIGDG 1259

RESULT 20

US-11-044-111-23
; Sequence 23, Application US/11044111
; Publication No. US20050272362A1
; GENERAL INFORMATION:
; APPLICANT: Chiang, Wen
; APPLICANT: Strasburg, Gale
; APPLICANT: Linz, John
; TITLE OF INVENTION: Genetic Test for PSE-Susceptible Turkeys
; FILE REFERENCE: MSU-09308
; CURRENT APPLICATION NUMBER: US/11/044,111
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 23
; LENGTH: 3375
; TYPE: PRT
; ORGANISM: Meleagris gallopavo
US-11-044-111-23

Query Match 82.5%; Score 33; DB 7; Length 3375;
Best Local Similarity 85.7%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGIGDG 7
Db 2951 GGDGDD 2957

RESULT 21

US-10-467-657-3590
; Sequence 3590, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 3590
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-3590

```
Query Match      80.0%; Score 32; DB 6; Length 91;
Best Local Similarity 71.4%; Pred. No. 60;
Matches 5; Conservative 1; Mismatches 1; Indels 1; Gaps 0;

QY      1 GGIGDGG 7
Db      37 GGVGGGG 43

RESULT 22
US-10-467-657-6440
; Sequence 6440, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 6440
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-6440

Query Match      80.0%; Score 32; DB 6; Length 91;
Best Local Similarity 71.4%; Pred. No. 60;
Matches 5; Conservative 1; Mismatches 1; Indels 1; Gaps 0;

QY      1 GGIGDGG 7
Db      37 GGVGGGG 43

RESULT 23
US-10-714-887-110
; Sequence 110, Application US/10714887
; Publication No. US20060015972A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: HEARD, Jacqueline
; APPLICANT: RIECHMANN, Jose Luis
; APPLICANT: CREELMAN, Robert
; APPLICANT: RATCLIFFE, Oliver
; APPLICANT: CANALES, Roger
; APPLICANT: REPETTI, Peter
; APPLICANT: KUMIMOTO, Roderick W
; APPLICANT: GUTTERSON, Neal
; APPLICANT: REUBER, T. Lynne
; APPLICANT: PINEDA, Omaira
; APPLICANT: SHERMAN, Bradley K
; TITLE OF INVENTION: PLANT TRANSCRIPTIONAL REGULATORS OF DROUGHT STRESS
; FILE REFERENCE: MBI0058-CIP
; CURRENT APPLICATION NUMBER: US/10/714,887
; CURRENT FILING DATE: 2003-11-13
; PRIOR APPLICATION NUMBER: 10/412,699
; PRIOR FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: 09/506,720
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/135,134
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: 09/394,519
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: 09/533,392
; PRIOR FILING DATE: 2000-03-22

; PRIOR APPLICATION NUMBER: 09/533,029
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/532,591
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,030
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/125,814
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: 09/713,994
; PRIOR FILING DATE: 2000-11-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 430
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 110
; LENGTH: 290
; TYPE: PRT
; ORGANISM: Oryza sativa (indica cultivar-group)
; FEATURE:
; OTHER INFORMATION: G3674 polypeptide Orthologous to G2999
US-10-714-887-110

Query Match      80.0%; Score 32; DB 6; Length 290;
Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 GGIGDGG 7
Db      44 GGVGGGG 50

RESULT 24
US-11-052-554A-167
; Sequence 167, Application US/11052554A
; Publication No. US2005028866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 167
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis H37Rv
US-11-052-554A-167

Query Match      80.0%; Score 32; DB 7; Length 434;
Best Local Similarity 85.7%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GGIGDGG 7
Db      153 GGGDGG 159

RESULT 25
US-11-059-814-20
; Sequence 20, Application US/11059814
; Publication No. US20050272650A1
; GENERAL INFORMATION:
; APPLICANT: University of South Florida
; APPLICANT: Mohapatra, Shyam
; TITLE OF INVENTION: Materials and Methods for Treatment of Inflammatory and Cell
; TITLE OF INVENTION: Proliferation Disorders
; FILE REFERENCE: USF-215PTCZ
; CURRENT APPLICATION NUMBER: US/11/059,814
```

; CURRENT FILING DATE: 2005-02-17
; PRIOR APPLICATION NUMBER: 60/521,072
; PRIOR FILING DATE: 2004-02-17
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20
; LENGTH: 541
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-059-814-20

Query Match 80.0%; Score 32; DB 7; Length 541;
Best Local Similarity 71.4%; Pred. No. 3.2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
||| ||

Db 27 GGVGGG 33

RESULT 26

US-10-504-364-3
; Sequence 3, Application US/10504364
; Publication No. US20050287158A1
; GENERAL INFORMATION:
; APPLICANT: Georgetown University
; APPLICANT: Jung, Mira
; TITLE OF INVENTION: The p65 subunit of NF-kB for the radiosensitization of cells
; FILE REFERENCE: 082137-0310990
; CURRENT APPLICATION NUMBER: US/10/504,364
; CURRENT FILING DATE: 2004-08-12
; PRIOR APPLICATION NUMBER: PCT/US03/04513
; PRIOR FILING DATE: 2003-02-19
; PRIOR APPLICATION NUMBER: US 60/356,748
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3
; LENGTH: 558
; TYPE: PRT
; ORGANISM: Gallus gallus
US-10-504-364-3

Query Match 80.0%; Score 32; DB 6; Length 558;
Best Local Similarity 85.7%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
||| ||

Db 454 GGIGAGG 460

RESULT 27

US-10-504-364-4
; Sequence 4, Application US/10504364
; Publication No. US20050287158A1
; GENERAL INFORMATION:
; APPLICANT: Georgetown University
; APPLICANT: Jung, Mira
; TITLE OF INVENTION: The p65 subunit of NF-kB for the radiosensitization of cells
; FILE REFERENCE: 082137-0310990
; CURRENT APPLICATION NUMBER: US/10/504,364
; CURRENT FILING DATE: 2004-08-12
; PRIOR APPLICATION NUMBER: PCT/US03/04513
; PRIOR FILING DATE: 2003-02-19
; PRIOR APPLICATION NUMBER: US 60/356,748
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4
; LENGTH: 558
; TYPE: PRT
; ORGANISM: Gallus gallus

US-10-504-364-4

Query Match 80.0%; Score 32; DB 6; Length 558;
Best Local Similarity 85.7%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
||| ||

Db 454 GGIGAGG 460

RESULT 28

US-11-052-554A-150
; Sequence 150, Application US/11052554A
; Publication No. US2005028866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 150
; LENGTH: 618
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis H37Rv
US-11-052-554A-150

Query Match 80.0%; Score 32; DB 7; Length 618;
Best Local Similarity 85.7%; Pred. No. 3.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
||| ||

Db 544 GGIGAGG 550

RESULT 29

US-11-045-802-22
; Sequence 22, Application US/11045802
; Publication No. US20050257289A1
; GENERAL INFORMATION:
; APPLICANT: Gordon-Kamm, William
; APPLICANT: Helentjaris, Tim
; APPLICANT: Lowe, Keith
; APPLICANT: Shen, Bo
; APPLICANT: Tarczynski, Mitchell
; APPLICANT: Zheng, Peizhong
; TITLE OF INVENTION: AP2 Domain Transcription Factor ODP2 (Ovule Development Protein
; FILE REFERENCE: 035718/286074
; CURRENT APPLICATION NUMBER: US/11/045,802
; CURRENT FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: 60/541,122
; PRIOR FILING DATE: 2004-02-02
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 710
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Variant of Zm-ODP2 having 87.3% amino acid
; OTHER INFORMATION: sequence identity to SEQ ID NO:2 (Zm-ODP2).
US-11-045-802-22

Query Match 80.0%; Score 32; DB 7; Length 710;

Best Local Similarity 85.7%; Pred. No. 4.1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
||| |||
Db 128 GGIHDGG 134

RESULT 30

US-11-045-802-23
; Sequence 23, Application US/11045802
; Publication No. US20050257289A1
; GENERAL INFORMATION:
; APPLICANT: Gordon-Kamm, William
; APPLICANT: Helentjaris, Tim
; APPLICANT: Lowe, Keith
; APPLICANT: Shen, Bo
; APPLICANT: Tarczynski, Mitchell
; APPLICANT: Zheng, Peizhong
; TITLE OF INVENTION: AP2 Domain Transcription Factor ODP2 (Ovule Development Protein 2)
; FILE REFERENCE: 035718/286074
; CURRENT APPLICATION NUMBER: US/11/045,802
; CURRENT FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: 60/541,122
; PRIOR FILING DATE: 2004-02-02
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 710
; TYPE: PRT
; ORGANISM: Artificial Sequence

; OTHER INFORMATION: Variant of Zm-ODP2 having 82.4% amino acid
; OTHER INFORMATION: sequence identity to SEQ ID NO:2 (Zm-ODP2).
US-11-045-802-23

Query Match 80.0%; Score 32; DB 7; Length 710;
Best Local Similarity 85.7%; Pred. No. 4.1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
||| |||
Db 128 GGIHDGG 134

RESULT 31

US-11-052-554A-154
; Sequence 154, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 154
; LENGTH: 767
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis H37Rv

Query Match 80.0%; Score 32; DB 7; Length 767;
Best Local Similarity 71.4%; Pred. No. 4.4e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
||| |||
Db 608 GGVGGGG 614

RESULT 32

US-11-205-109-18
; Sequence 18, Application US/11205109
; Publication No. US20050287641A1
; GENERAL INFORMATION:
; APPLICANT: Farnet, Chris
; APPLICANT: Zazopoulos, Emmanuel
; APPLICANT: Staffa, Alfredo
; TITLE OF INVENTION: GENE CLUSTER FOR RAMOPLANIN BIOSYNTHESIS
; FILE REFERENCE: 3002-2US
; CURRENT APPLICATION NUMBER: US/11/205,109
; CURRENT FILING DATE: 2005-08-17
; PRIOR APPLICATION NUMBER: US/09/976,059
; PRIOR FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: US 60/239,924
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
; LENGTH: 891
; TYPE: PRT
; ORGANISM: Actinoplanes sp.
; NAME/KEY: misc.feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: V is a non-standard initiator codon. It is expected that the bi
; OTHER INFORMATION: synthesized protein will have a formylmethionine residue at this
; OTHER INFORMATION: position
US-11-205-109-18

Query Match 80.0%; Score 32; DB 7; Length 891;
Best Local Similarity 83.3%; Pred. No. 5.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 GGIGDGG 7
||| |||
Db 134 GLGDGG 139

RESULT 33

US-11-052-554A-147
; Sequence 147, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 147
; LENGTH: 923
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis H37Rv

Query Match 80.0%; Score 32; DB 7; Length 923;
Best Local Similarity 71.4%; Pred. No. 5.2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
||| |||

Db 352 GGVGGG 358

RESULT 34

US-10-523-503-44
; Sequence 44, Application US/10523503
; Publication No. US20060037102A1
; GENERAL INFORMATION:
; APPLICANT: BASF PLANT SCIENCE GMBH
; TITLE OF INVENTION: SUGAR AND LIPID METABOLISM REGULATORS IN PLANTS IV
; FILE REFERENCE: 16313-0236
; CURRENT APPLICATION NUMBER: US/10/523,503
; CURRENT FILING DATE: 2005-02-02
; PRIOR APPLICATION NUMBER: US 60/400,803
; PRIOR FILING DATE: 2002-08-02
; NUMBER OF SEQ ID NOS: 163
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 44
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-523-503-44

Query Match 77.5%; Score 31; DB 6; Length 212;

Best Local Similarity 85.7%; Pred. No. 1.9e+02;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGVGGG 7

Db 189 GGVGG 195

RESULT 35

US-10-714-887-232
; Sequence 232, Application US/10714887
; Publication No. US20060015972A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: HEARD, Jacqueline
; APPLICANT: RIECHMANN, Jose Luis
; APPLICANT: CREELMAN, Robert
; APPLICANT: RATCLIFFE, Oliver
; APPLICANT: CANALES, Roger
; APPLICANT: REPETTI, Peter
; APPLICANT: KUMIMOTO, Roderick W
; APPLICANT: GUTTERSON, Neal
; APPLICANT: REUBER, T. Lynne
; APPLICANT: PINEDA, Omalra
; APPLICANT: SHERMAN, Bradley K
; TITLE OF INVENTION: PLANT TRANSCRIPTIONAL REGULATORS OF DROUGHT STRESS
; FILE REFERENCE: MB10058-CIP
; CURRENT APPLICATION NUMBER: US/10/714,887
; CURRENT FILING DATE: 2003-11-13
; PRIOR APPLICATION NUMBER: 10/412,699
; PRIOR FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: 09/506,720
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/135,134
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: 09/394,519
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: 09/533,392
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,029
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/532,591
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,030
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/125,814
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: 09/713,994
; PRIOR FILING DATE: 2000-11-16

; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 430
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 232
; LENGTH: 265
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G634 polypeptide reference sequence
US-10-714-887-232

Query Match 77.5%; Score 31; DB 6; Length 265;

Best Local Similarity 71.4%; Pred. No. 2.3e+02;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGVGGG 7

Db 41 GGVGGG 47

RESULT 36

US-10-467-657-5426
; Sequence 5426, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 5426
; LENGTH: 270
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-5426

Query Match 77.5%; Score 31; DB 6; Length 270;

Best Local Similarity 71.4%; Pred. No. 2.4e+02;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGVGGG 7

Db 157 GGVGGG 163

RESULT 37

US-11-124-367A-460
; Sequence 460, Application US/11124367A
; Publication No. US20060024700A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: Hongjin Huang
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; TITLE OF INVENTION: Fibrosis Methods of Detection and Uses Thereof
; FILE REFERENCE: CL001519.ORD
; CURRENT APPLICATION NUMBER: US/11/124,367A
; CURRENT FILING DATE: 2005-05-09
; PRIOR APPLICATION NUMBER: US 60/568,846
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/582,609
; PRIOR FILING DATE: 2004-06-25
; PRIOR APPLICATION NUMBER: US 60/599,554
; PRIOR FILING DATE: 2004-08-09
; NUMBER OF SEQ ID NOS: 34460
; SOFTWARE: FastSeq for Windows Version 4.0

```
; SEQ ID NO 460
; LENGTH: 430
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-124-367A-460

Query Match          77.5%; Score 31; DB 7; Length 430;
Best Local Similarity 71.4%; Pred. No. 3.7e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 GGIGDGG 7
Db      56 GGMGSGG 62

RESULT 38
US-11-110-082-32
; Sequence 32, Application US/11110082
; Publication No. US20050266558A1
; GENERAL INFORMATION:
; APPLICANT: Demmer, Jeroen
; APPLICANT: Hall, Claire
; APPLICANT: Norriss, Michael Geoffrey
; APPLICANT: Saulsbury, Keith Martin
; APPLICANT: Saulsbury, Keith Martin
; TITLE OF INVENTION: Compositions Isolated from Forage
; TITLE OF INVENTION: Grasses and methods for their use.
; FILE REFERENCE: 11000.10740c1
; CURRENT APPLICATION NUMBER: US/11/110,082
; CURRENT FILING DATE: 2005-04-19
; PRIOR APPLICATION NUMBER: 60/563,723
; PRIOR FILING DATE: 2004-04-20
; PRIOR APPLICATION NUMBER: 10/655,799
; PRIOR FILING DATE: 2003-09-05
; PRIOR APPLICATION NUMBER: 60/408,782
; PRIOR FILING DATE: 2002-09-05
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 530
; TYPE: PRT
; ORGANISM: Lolium perenne
US-11-110-082-32

Query Match          77.5%; Score 31; DB 7; Length 530;
Best Local Similarity 71.4%; Pred. No. 4.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 GGIGDGG 7
Db      435 GGAGEGG 441

RESULT 39
US-11-072-175-243
; Sequence 243, Application US/11072175
; Publication No. US20060029944A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT
; TITLE OF INVENTION: INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
; TITLE OF INVENTION: PROTEIN TYROSINE KINASE PATHWAYS IN BREAST CELLS
; FILE REFERENCE: D0273A CIP
; CURRENT APPLICATION NUMBER: US/11/072,175
; CURRENT FILING DATE: 2005-03-05
; PRIOR APPLICATION NUMBER: US 60/406,385
; PRIOR FILING DATE: 2002-08-27
; PRIOR APPLICATION NUMBER: US 10/648,593
; PRIOR FILING DATE: 2003-08-26
; NUMBER OF SEQ ID NOS: 571
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 243
; LENGTH: 547
; TYPE: PRT

; ORGANISM: Homo sapiens
US-11-072-175-243

Query Match          77.5%; Score 31; DB 7; Length 547;
Best Local Similarity 85.7%; Pred. No. 4.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GGIGDGG 7
Db      326 GGIGFGG 332

RESULT 40
US-11-072-512-3324
; Sequence 3324, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOUYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: Novel full length cdNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3324
; LENGTH: 602
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-072-512-3324

Query Match          77.5%; Score 31; DB 7; Length 602;
Best Local Similarity 71.4%; Pred. No. 5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 GGIGDGG 7
Db      72 GGLGPGG 78

RESULT 41
US-11-052-554A-163
; Sequence 163, Application US/11052554A
; Publication No. US2005028866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
```

1
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 163
; LENGTH: 606
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis H37RV
US-11-052-554A-163

Query Match 77.5%; Score 31; DB 7; Length 606;
Best Local Similarity 71.4%; Pred. No. 5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
||:|
Db 560 GGVGAGG 566

RESULT 42

US-10-661-966-14
; Sequence 14, Application US/10661966
; Publication No. US20050277118A1
; GENERAL INFORMATION:
; APPLICANT: Roth, Richard B.
; APPLICANT: Nelson, Matthew Roberts
; APPLICANT: Braun, Andreas
; TITLE OF INVENTION: METHODS FOR IDENTIFYING SUBJECTS AT RISK
; FILE REFERENCE: 524592003800
; CURRENT APPLICATION NUMBER: US/10/661,966
; PRIOR FILING DATE: 2003-09-11
; PRIOR APPLICATION NUMBER: 60/410,595
; PRIOR FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: 60/422,344
; PRIOR FILING DATE: 2002-10-29
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 659
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-661-966-14

Query Match 77.5%; Score 31; DB 6; Length 659;
Best Local Similarity 71.4%; Pred. No. 5.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
||:|
Db 3 GCGGEGG 9

RESULT 43

US-11-052-554A-138
; Sequence 138, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 138
; LENGTH: 1381
; TYPE: PRT

; ORGANISM: Mycobacterium tuberculosis H37RV
US-11-052-554A-138

Query Match 77.5%; Score 31; DB 7; Length 1381;
Best Local Similarity 71.4%; Pred. No. 1.1e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
||:|
Db 399 GGVGAGG 405

RESULT 44

US-11-052-554A-135
; Sequence 135, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 135
; LENGTH: 1901
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis H37RV
US-11-052-554A-135

Query Match 77.5%; Score 31; DB 7; Length 1901;
Best Local Similarity 71.4%; Pred. No. 1.5e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
||:|
Db 1785 GGLGGGG 1791

RESULT 45

US-10-939-890-492
; Sequence 492, Application US/10939890
; Publication No. US20050250700A1
; GENERAL INFORMATION:
; APPLICANT: Sato, Aaron K.
; APPLICANT: Sexton, Daniel J.
; APPLICANT: Dransfield, Daniel T.
; APPLICANT: Ladner, Robert C.
; APPLICANT: Arbogast, Christophe
; APPLICANT: Bussat, Philippe
; APPLICANT: Fan, Hong
; APPLICANT: Khurana, Sudha
; APPLICANT: Linder, Karen E.
; APPLICANT: Marinelli, Edmund R.
; APPLICANT: Nanjappan, Palaniappa
; APPLICANT: Nunn, Adrian D.
; APPLICANT: Pillai, Radhakrishna
; APPLICANT: Pochon, Sibylle
; APPLICANT: Ramalingam, Kondareddiar
; APPLICANT: Shrivastava, Ajay
; APPLICANT: Song, Bo
; APPLICANT: Swenson, Rolf E.
; APPLICANT: Von Wronski, Mathew A.
; TITLE OF INVENTION: KDR AND VEGF/KDR BINDING PEPTIDES
; FILE REFERENCE: D0617.70014US00
; CURRENT APPLICATION NUMBER: US/10/939,890
; CURRENT FILING DATE: 2004-09-13
; PRIOR APPLICATION NUMBER: US 10/661,156

; PRIOR FILING DATE: 2003-09-11
; PRIOR APPLICATION NUMBER: US 10/382,082
; PRIOR FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: PCT/US03/06731
; PRIOR FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: US 60/440,411
; PRIOR FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: US 60/360,851
; PRIOR FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 883
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 492
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Polypeptide Linker
US-10-939-890-492

Query Match 75.0%; Score 30; DB 6; Length 16;
Best Local Similarity 71.4%; Pred. No. 24;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
||| |||
Db 10 GSGEGG 16

RESULT 46
US-10-467-657-922
; Sequence 922, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 922
; LENGTH: 189
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-922

Query Match 75.0%; Score 30; DB 6; Length 189;
Best Local Similarity 83.3%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDG 6
||| |||
Db 170 GGFGDG 175

RESULT 47
US-10-467-657-2106
; Sequence 2106, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:

; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 2106
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-2106

Query Match 75.0%; Score 30; DB 6; Length 200;
Best Local Similarity 71.4%; Pred. No. 2.6e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
||| |||
Db 143 GGVGVGG 149

RESULT 48
US-10-883-512-88
; Sequence 88, Application US/10883512
; Publication No. US20060005265A1
; GENERAL INFORMATION:
; APPLICANT: Bughrara, Suleiman
; APPLICANT: Han, Zhao
; APPLICANT: Wang, Yuexia
; TITLE OF INVENTION: RyeGrass CBF3 Gene: Identification and Isolation
; FILE REFERENCE: MSU-08807
; CURRENT APPLICATION NUMBER: US/10/883,512
; CURRENT FILING DATE: 2004-07-01
; NUMBER OF SEQ ID NOS: 199
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 88
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Cucumis melo
US-10-883-512-88

Query Match 75.0%; Score 30; DB 6; Length 229;
Best Local Similarity 71.4%; Pred. No. 2.9e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
||| |||
Db 167 GGVGVGG 173

RESULT 49
US-11-232-405A-40
; Sequence 40, Application US/11232405A
; Publication No. US20060035347A1
; GENERAL INFORMATION:
; APPLICANT: USUDA, Yoshihiro
; APPLICANT: NISHIO, Yosuke
; APPLICANT: YASUEDA, Hisashi
; APPLICANT: SUGIMOTO, Shinichi
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING POLYPEPTIDES INVOLVED IN AMINO ACID BIO
; TITLE OF INVENTION: IN METHYLOPHILUS METHYLOTROPUS
; FILE REFERENCE: 232743US0
; CURRENT APPLICATION NUMBER: US/11/232,405A
; CURRENT FILING DATE: 2005-09-22
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 40
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Methylophilus methylotrophus
US-11-232-405A-40

Query Match 75.0%; Score 30; DB 7; Length 275;

Best Local Similarity 71.4%; Pred. No. 3.4e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGIGDGG 7
|||:|
Db 162 GGVGIGG 168

RESULT 50

US-10-714-887-300
; Sequence 300, Application US/10714887
; Publication No. US20060015972A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: HEARD, Jacqueline
; APPLICANT: RIECHMANN, Jose Luis
; APPLICANT: CREELMAN, Robert
; APPLICANT: RATCLIFFE, Oliver
; APPLICANT: CANALES, Roger
; APPLICANT: REPETTI, Peter
; APPLICANT: KUMIMOTO, Roderick W
; APPLICANT: GUTTERSON, Neal
; APPLICANT: REUBER, T. Lynne
; APPLICANT: PINEDA, Omaira
; APPLICANT: SHERMAN, Bradley K
; TITLE OF INVENTION: PLANT TRANSCRIPTIONAL REGULATORS OF DROUGHT STRESS
; FILE REFERENCE: MB10058-CIP
; CURRENT APPLICATION NUMBER: US/10/714,887
; CURRENT FILING DATE: 2003-11-13
; PRIOR APPLICATION NUMBER: 10/412,699
; PRIOR FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: 09/506,720
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/135,134
; PRIOR FILING DATE: 1998-05-20
; PRIOR APPLICATION NUMBER: 09/394,519
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: 09/533,392
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,029
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/532,591
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,030
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/125,814
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: 09/713,994
; PRIOR FILING DATE: 2000-11-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 430
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 300
; LENGTH: 324
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G1070 polypeptide Functionally related and homologous to G1073
US-10-714-887-300

Query Match 75.0%; Score 30; DB 6; Length 324;
Best Local Similarity 71.4%; Pred. No. 4e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGIGDGG 7
|||:|
Db 88 GGSGEgg 94

Search completed: February 27, 2006, 10:04:55
Job time : 18 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 27, 2006, 09:46:09 ; Search time 39 Seconds
(without alignments)
17.270 Million cell updates/sec

Title: US-10-027-015b-2

Perfect score: 40

Sequence: 1 GGIGDGG 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

PIR 80:**

1: pirl:**

2: pirl:**

3: pirl:**

4: pirl:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	256	2 F75310	hypothetical prote
2	40	100.0	289	2 T14694	DNA polymerase I -
3	40	100.0	331	2 T15026	hypothetical prote
4	40	100.0	351	2 S70765	nodulin-45 precurs
5	40	100.0	425	2 G83299	probable outer mem
6	40	100.0	1090	2 A41696	regulatory protein
7	36	90.0	223	2 B96506	hypothetical prote
8	36	90.0	224	2 G87388	hypothetical prote
9	36	90.0	368	2 T51111	aminotransferase I
10	36	90.0	526	1 KRBOVI	keratin, 54K type
11	36	90.0	628	2 A69381	type II secretion
12	36	90.0	1891	2 T43262	calcium channel al
13	35	87.5	117	2 D89075	protein K04A8.3 li
14	35	87.5	185	2 D84538	probable glycine-r
15	35	87.5	220	2 A44805	eggshell protein p
16	35	87.5	304	2 AE1069	mrr restriction sy
17	35	87.5	304	2 A40368	mrr protein - Esch
18	35	87.5	317	2 T20018	hypothetical prote
19	35	87.5	366	2 T36689	probable transp
20	35	87.5	382	2 T04347	teosinte branched1
21	35	87.5	465	1 S01820	glycine-rich cell
22	35	87.5	596	1 WMBEH2	ULJ2 protein - hum
23	35	87.5	718	2 D84675	hypothetical prote
24	35	87.5	719	2 T52510	hypothetical prote
25	35	87.5	1079	2 B70807	hypothetical glyci
26	34	85.0	234	2 F95952	probable membrane
27	34	85.0	266	2 G71208	hypothetical prote
28	34	85.0	268	2 H75021	hypothetical prote
29	34	85.0	310	2 H97338	dioxygenase relate

30	34	85.0	325	2 G83979	hypothetical prote
31	34	85.0	352	1 CEECFX	cell division prot
32	34	85.0	415	2 T46716	hypothetical prote
33	34	85.0	430	2 T12583	oleoyl-[acyl-carri
34	34	85.0	431	2 T46437	hypothetical prote
35	34	85.0	476	2 AC2306	hypothetical prote
36	34	85.0	495	2 T17478	hypothetical prote
37	34	85.0	496	2 H70839	hypothetical glyci
38	34	85.0	507	2 T33024	hypothetical prote
39	34	85.0	508	2 B49131	ryanodine receptor
40	34	85.0	572	2 T40856	probable nucleotid
41	34	85.0	575	2 T48224	probable homeodoma
42	34	85.0	658	1 S23391	transcription fact
43	34	85.0	723	2 T14605	probable cell divi
44	34	85.0	731	2 C70974	hypothetical glyci
45	34	85.0	805	1 S19738	transitional endop
46	34	85.0	806	1 VPPG	transitional endop
47	34	85.0	806	1 A55190	transitional endop
48	34	85.0	806	1 S25197	transitional endop
49	34	85.0	806	2 T02243	probable transitio
50	34	85.0	822	2 S70012	methionine-S-oxide
51	34	85.0	1039	2 C84745	hypothetical prote
52	34	85.0	1077	2 T38900	probable alpha-man
53	34	85.0	1327	2 T09402	immunoglobulin-lik
54	34	85.0	2561	2 T24864	hypothetical prote
55	34	85.0	5126	2 S40450	ryanodine receptor
56	33	82.5	100	2 S20482	hypothetical prote
57	33	82.5	103	2 F83918	hypothetical prote
58	33	82.5	123	2 T51568	hypothetical prote
59	33	82.5	159	2 T37098	hypothetical prote
60	33	82.5	169	2 T02081	ABA- and ripening-
61	33	82.5	180	2 T49530	related to glycine
62	33	82.5	192	2 T36038	probable export as
63	33	82.5	200	2 B86181	hypothetical prote
64	33	82.5	201	2 T00799	hypothetical prote
65	33	82.5	231	2 G84722	hypothetical prote
66	33	82.5	239	2 T01406	bovine corneal pro
67	33	82.5	247	2 T39104	hypothetical prote
68	33	82.5	258	2 S71561	drought-induced pr
69	33	82.5	260	2 F69337	thiamin biosynthes
70	33	82.5	267	2 G90006	hypothetical prote
71	33	82.5	288	2 T32295	hypothetical prote
72	33	82.5	291	2 T26791	hypothetical prote
73	33	82.5	305	2 A83061	hypothetical prote
74	33	82.5	308	2 E81288	hypothetical prote
75	33	82.5	320	2 T09555	hypothetical prote
76	33	82.5	326	2 H72674	fibillararin - Arab
77	33	82.5	342	2 A30189	hypothetical prote
78	33	82.5	342	2 S75086	iron stress-induce
79	33	82.5	367	2 JC4831	adsorption protein
80	33	82.5	389	2 T51355	membrane protein I
81	33	82.5	408	2 S57483	glycin-rich protei
82	33	82.5	410	2 T1960	hypothetical prote
83	33	82.5	431	2 T04956	hypothetical prote
84	33	82.5	440	2 T01768	hypothetical prote
85	33	82.5	444	2 T05614	hypothetical prote
86	33	82.5	446	2 S06938	glutamate dehydrog
87	33	82.5	453	2 A41640	vestigial protein
88	33	82.5	457	2 T51528	hypothetical prote
89	33	82.5	458	2 B96643	hypothetical prote
90	33	82.5	460	2 A12264	aldehyde dehydroge
91	33	82.5	462	2 T05995	hypothetical prote
92	33	82.5	476	2 T47941	hypothetical prote
93	33	82.5	498	2 T51430	dolichyl-phosphate
94	33	82.5	533	2 B86325	T29M8.4 protein -
95	33	82.5	557	2 AD2832	DNA repair protein
96	33	82.5	557	2 H97609	period clock prote
97	33	82.5	571	1 UMFF	keratin 1, type II
98	33	82.5	643	1 KKHU2	probable WD-40 rep
99	33	82.5	648	2 H84587	hypothetical glyci
100	33	82.5	714	2 A70807	

ALIGNMENTS

RESULT 1

F75310
hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C/Accession: F75310
R/White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A>Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A/Reference number: A75250; MUID:20036896; PMID:10567266
A/Accession: F75310
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-256 <WHI>
A/Cross-references: UNIPROT:Q9RSI7; UNIPARC:UPI00000C1A78; GB:AE002048; GB:AE000513; NID
A/Experimental source: strain R1
C/Genetics:
A/Gene: DR2137
A/Map position: 1

Query Match 100.0%; Score 40; DB 2; Length 256;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
|||||
DB 154 GGIGDGG 160

RESULT 2

T14694
DNA polymerase I - Yersinia pestis plasmid pMT1
C:Species: Yersinia pestis
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C/Accession: T14694
R/Hu, P.; Elliott, J.; McCready, P.; Skowronski, E.; Garnes, J.; Kobayashi, A.; Carrano,
submitted to the EMBL Data Library, March 1998
A/Description: Structural organization of virulence determinants in three Yersinia pestis
A/Reference number: Z18168
A/Accession: T14694
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-289 <HUP>
A/Cross-references: UNIPROT:O68765; UNIPROT:Q9RIC9; UNIPARC:UPI000016E2B2; EMBL:AF053947
C/Genetics:
A/Genome: plasmid pMT1

Query Match 100.0%; Score 40; DB 2; Length 289;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
|||||
DB 171 GGIGDGG 177

RESULT 3

T15026
hypothetical protein Y1111 - Yersinia pestis plasmid pMT1
C:Species: Yersinia pestis
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C/Accession: T15026
R/Lindler, L.E.; Plano, G.V.; Burland, V.; Mayhew, G.F.; Blattner, F.R.
Infect. Immun. 66, 5731-5742, 1998
A>Title: Complete DNA sequence and detailed analysis of the Yersinia pestis KIMS plasmid
A/Reference number: Z18268; MUID:99043898; PMID:9826348
A/Accession: T15026
A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA
A/Residues: 1-331 <LIN>
A/Cross-references: UNIPROT:O68765; UNIPARC:UPI00000B4994; EMBL:AF074611; NID:93883003;
C/Genetics:
A/Gene: Y1111
A/Genome: plasmid pMT1

Query Match 100.0%; Score 40; DB 2; Length 331;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
|||||
DB 213 GGIGDGG 219

RESULT 4

S70765
nodulin-45 precursor - narrow-leaved blue lupine
C:Species: Lupinus angustifolius (narrow-leaved blue lupine)
C>Date: 15-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 09-Jul-2004
C/Accession: S70765; S34261; S34260
R/Rice, S.J.; Grant, M.R.; Reynolds, P.H.S.; Farnden, K.J.F.
Plant Sci. 90, 155-166, 1993
A>Title: DNA sequence of nodulin-45 from Lupinus angustifolius.
A/Reference number: S70765
A/Accession: S70765
A/Molecule type: DNA
A/Residues: 1-63 <RIC>
A/Cross-references: UNIPROT:Q43494; UNIPROT:Q43878; UNIPARC:UPI000017B054; EMBL:Z12820
A/Accession: S34261
A/Molecule type: mRNA
A/Residues: 10-20, 'R', 22-351 <RIW>
A/Cross-references: UNIPARC:UPI000017B055; EMBL:Z12625; NID:9311929; PID:9311930
C/Genetics:
A/Introns: 46/1
A/Note: the list of introns may be incomplete
C/Keywords: nodulation
F/-24/Domain: signal sequence #status predicted <SIG>
F/25-351/Product: nodulin-45 #status predicted <MAR>

Query Match 100.0%; Score 40; DB 2; Length 351;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
|||||
DB 272 GGIGDGG 278

RESULT 5

G83299
probable outer membrane protein PA2760 [imported] - Pseudomonas aeruginosa (strain PAO1)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C/Accession: G83299
R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lin
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A>Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic path
A/Reference number: A82950; MUID:20437337; PMID:10984043
A/Accession: G83299
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-425 <STO>
A/Cross-references: UNIPROT:Q9I083; UNIPARC:UPI00000C57ED; GB:AE004704; GB:AE004091; NI
C/Genetics:
A/Experimental source: strain PAO1

Query Match 100.0%; Score 40; DB 2; Length 425;

C>Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 09-Jul-2004
C:Accession: A02941; S07262
R:Rieger, W.; Jorcano, J.L.; Franke, W.W.
EMBO J. 4, 2261-2267, 1985
A:Title: Complete sequence of a bovine type I cytokeratin gene: conserved and variable regions
A:Reference number: A02941; MUID:86081734; PMID:2416562
A:Accession: A02941
A:Molecule type: DNA
A:Residues: 1-526 <RIE>
A:Cross-references: UNIPROT:P06394; UNIPARC:UPI000012DAE2; GB:X02870; NID:9478; PIDN:CAP
R:Jorcano, J.L.; Rieger, M.; Franz, J.K.; Schiller, D.L.; Moll, R.; Franke, W.W.
J. Mol. Biol. 179, 257-281, 1984
A:Title: Identification of two types of keratin polypeptides within the acidic cytokeratin
A:Reference number: S07262; MUID:85058191; PMID:6209405
A:Accession: S07262
A:Molecule type: mRNA
A:Residues: 281-466; 'PAAATAAEVQRWRPRXFRQYQ', 490-491, 494, 'PVAVARRRK', 504-526 <JOR>
A:Cross-references: UNIPARC:UPI0000173D56; EMBL:X01460
A:Note: this sequence has been revised in reference A02941
C:Genetics:
A:Gene: cytokeratin VIb
A:Introns: 190/3; 218/2; 270/3; 324/3; 366/3; 439/2; 525/2
C:Superfamily: cytoskeletal keratin
C:Keywords: coiled coil; intermediate filament
F:1-127/Domain: head <HD>
F:1-127/Region: EI and VI subdomains
F:128-441/Domain: rod <ROD>
F:128-162/Region: coil 1A
F:163-176/Region: linker 1
F:177-277/Region: coil 1B
F:278-293/Region: linker 12
F:294-312/Region: coil 2A
F:313-320/Region: linker 2
F:321-441/Region: coil 2B
F:379/Region: stutter
F:442-526/Domain: tail <END>
F:442-526/Region: V2 and E2 subdomains
Query Match 90.0%; Score 36; DB 1; Length 526;
Best Local Similarity 85.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGIGDGG 7
DB 115 GGFDDGG 121
RESULT 11
A69381
type II secretion system protein (gspE-4) homolog - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: A69381
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.P.
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaea
A:Reference number: A69250; MUID:98049343; PMID:9389475
A:Accession: A69381
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-628 <LE>
A:Cross-references: UNIPROT:Q29133; UNIPARC:UPI0000056E8D; GB:AE001030; GB:AE000782; NID
C:Superfamily: conserved hypothetical protein M70900
Query Match 90.0%; Score 36; DB 2; Length 628;
Best Local Similarity 85.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGIGDGG 7

C>Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 09-Jul-2004
C:Accession: A02941; S07262
R:Rieger, W.; Jorcano, J.L.; Franke, W.W.
EMBO J. 4, 2261-2267, 1985
A:Title: Complete sequence of a bovine type I cytokeratin gene: conserved and variable regions
A:Reference number: A02941; MUID:86081734; PMID:2416562
A:Accession: A02941
A:Molecule type: DNA
A:Residues: 1-526 <RIE>
A:Cross-references: UNIPROT:P06394; UNIPARC:UPI000012DAE2; GB:X02870; NID:9478; PIDN:CAP
R:Jorcano, J.L.; Rieger, M.; Franz, J.K.; Schiller, D.L.; Moll, R.; Franke, W.W.
J. Mol. Biol. 179, 257-281, 1984
A:Title: Identification of two types of keratin polypeptides within the acidic cytokeratin
A:Reference number: S07262; MUID:85058191; PMID:6209405
A:Accession: S07262
A:Molecule type: mRNA
A:Residues: 281-466; 'PAAATAAEVQRWRPRXFRQYQ', 490-491, 494, 'PVAVARRRK', 504-526 <JOR>
A:Cross-references: UNIPARC:UPI0000173D56; EMBL:X01460
A:Note: this sequence has been revised in reference A02941
C:Genetics:
A:Gene: cytokeratin VIb
A:Introns: 190/3; 218/2; 270/3; 324/3; 366/3; 439/2; 525/2
C:Superfamily: cytoskeletal keratin
C:Keywords: coiled coil; intermediate filament
F:1-127/Domain: head <HD>
F:1-127/Region: EI and VI subdomains
F:128-441/Domain: rod <ROD>
F:128-162/Region: coil 1A
F:163-176/Region: linker 1
F:177-277/Region: coil 1B
F:278-293/Region: linker 12
F:294-312/Region: coil 2A
F:313-320/Region: linker 2
F:321-441/Region: coil 2B
F:379/Region: stutter
F:442-526/Domain: tail <END>
F:442-526/Region: V2 and E2 subdomains
Query Match 90.0%; Score 36; DB 1; Length 526;
Best Local Similarity 85.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGIGDGG 7
DB 115 GGFDDGG 121
RESULT 11
A69381
type II secretion system protein (gspE-4) homolog - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: A69381
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.P.
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaea
A:Reference number: A69250; MUID:98049343; PMID:9389475
A:Accession: A69381
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-628 <LE>
A:Cross-references: UNIPROT:Q29133; UNIPARC:UPI0000056E8D; GB:AE001030; GB:AE000782; NID
C:Superfamily: conserved hypothetical protein M70900
Query Match 90.0%; Score 36; DB 2; Length 628;
Best Local Similarity 85.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGIGDGG 7

Db 425 GGIGEGG 431
RESULT 12
T43262
calcium channel alpha-1 chain, L-type - Stylophora pistillata
C:Species: Stylophora pistillata
C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C:Accession: T43262
R:Zoccola, D.; Tambutte, E.; Senegas-Balas, F.; Michiels, J.F.; Failla, J.P.; Jaubert, A.
Gene 227, 157-167, 1999
A:Title: Cloning of a calcium channel alpha1 subunit from the reef-building coral, Styl
A:Reference number: Z22375; MUID:99148007; PMID:10023047
A:Accession: T43262
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1891 <ZOC>
A:Cross-references: UNIPROT:O97017; UNIPARC:UPI0000074851; EMBL:U64465; NID:94204977; P
C:Genetics:
A:Gene: CACHL
C:Superfamily: voltage-dependent calcium channel protein alpha-1 chain
Query Match 90.0%; Score 36; DB 2; Length 1891;
Best Local Similarity 85.7%; Pred. No. 4.6e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGIGDGG 7
DB 665 GGIGEGG 671
RESULT 13
D89075
protein K04A8.3 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: D89075
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolo
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/projects/C.el
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; an
A:Accession: D89075
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-117 <STO>
A:Cross-references: UNIPROT:Q94257; UNIPARC:UPI0000079DB2; GB:chr_V; PIDN:AAC48049.1; P
C:Genetics:
A:Gene: K04A8.3
A:Map position: 5
Query Match 87.5%; Score 35; DB 2; Length 117;
Best Local Similarity 71.4%; Pred. No. 40;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGIGDGG 7
DB 42 GGVGEGG 48
RESULT 14
D84538
probable glycine-rich RNA-binding protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana [mouse-ear cress]
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: D84538
R:Jin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: D84538
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-185 <STO>
A:Cross-references: UNIPROT:Q9SIX3; UNIPARC:UPI00000A035C; GB:AE002093; NID:g4544401; PID:10617197
C:Genetics:
A:Gene: At2g16260
A:Map position: 2
C:Superfamily: glycine-rich RNA-binding protein; ribonucleoprotein repeat homology

Query Match 87.5%; Score 35; DB 2; Length 185;
Best Local Similarity 85.7%; Pred. No. 63;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGI G D G G 7
DB 165 G G Y G D G G 171

RESULT 15
A44805
eggshell protein precursor - fluke (Schistosoma haematobium)
C:Species: Schistosoma haematobium
C:Date: 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 31-Dec-2004
A:Accession: A44805; C44805
R:Bobek, L.A.; LoVerde, P.T.; Rekosh, D.M.
Exp. Parasitol. 68, 17-30, 1989
A:Title: Schistosoma haematobium: analysis of eggshell protein genes and their expression
A:Reference number: A44805; MUID:89137380; PMID:2917627
A:Accession: A44805
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-220 <BOB>
A:Cross-references: UNIPROT:O96853; UNIPARC:UPI0000177E50; GB:M27659
A:Accession: C44805
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-70, 96-140, 142-220 <BO2>
A:Cross-references: UNIPARC:UPI000007BBA6; GB:M27658; NID:g160978; PID:g160979
C:Keywords: egg shell
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-220/Product: eggshell protein #status predicted <MAT>

Query Match 87.5%; Score 35; DB 2; Length 220;
Best Local Similarity 85.7%; Pred. No. 76;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGI G D G G 7
DB 40 G G Y G D G G 46

RESULT 16
AE1069
mrr restriction system protein [imported] - Salmonella enterica subsp. enterica serovar
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AE1069
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AE1069
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-304 <PAR>
A:Cross-references: UNIPARC:UPI000005A9C3; GB:AL513382; PIDN:CAD03372.1; PID:g16505642;
C:Genetics:

A:Gene: mrr

Query Match 87.5%; Score 35; DB 2; Length 304;
Best Local Similarity 85.7%; Pred. No. 11e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGI G D G G 7
DB 195 G G T G D G G 201

RESULT 17

A40368

mrr protein - Escherichia coli (strain K-12)

C:Species: Escherichia coli

C:Date: 21-Feb-1992 #sequence_revision 21-Feb-1992 #text_change 09-Jul-2004

C:Accession: A40368; S56577; A65250

R:Waite-Rees, P.A.; Keating, C.J.; Moran, L.S.; Slatko, B.E.; Hornstra, L.J.; Benner, J.

J. Bacteriol. 173, 5207-5219, 1991

A:Title: Characterization and expression of the Escherichia coli Mrr restriction system

A:Reference number: A40368; MUID:91317743; PMID:1650347

A:Accession: A40368

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-304 <WAI>

A:Cross-references: UNIPROT:P24202; UNIPARC:UPI000012F589; GB:X54198; NID:g42018; PIDN:1650347

R:Burland, V.; Plunkett III, G.; Sofia, H.J.; Daniels, D.L.; Blattner, F.R.

Nucleic Acids Res. 23, 2105-2119, 1995

A:Title: Analysis of the Escherichia coli genome VI: DNA sequence of the region from 9

A:Reference number: S56314; MUID:95334362; PMID:7610040

A:Accession: S56577

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-304 <BUR>

A:Cross-references: UNIPARC:UPI000012F589; EMBL:U14003; NID:g1263172; PIDN:AAA97248.1;

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1994

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;

A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617; PMID:9278503

A:Accession: A65250

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-304 <BLAT>

A:Cross-references: UNIPARC:UPI000012F589; GB:AE000506; GB:U00096; NID:g2367377; PIDN:1650347

A:Experimental source: strain K-12, substrain MGI655

C:Genetics:

A:Gene: mrr

Query Match 87.5%; Score 35; DB 2; Length 304;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGI G D G G 7
DB 195 G G T G D G G 201

RESULT 18

T20018

hypothetical protein C47B8.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T20018

R:Wild, A.

submitted to the EMBL Data Library, June 1996

A:Reference number: Z19211

A:Accession: T20018

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-317 <WIL>

A:Cross-references: UNIPROT:Q18687; UNIPARC:UPI0000079C20; EMBL:275530; PIDN:CAA99792.1

A;Experimental source: clone C47E8
C;Genetics:
A;Gene: CESP:C47E8.3
A;Map position: 5
A;Introns: 97/2; 194/3; 231/2

Query Match 87.5%; Score 35; DB 2; Length 317;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
||| |||||
Db 284 GGAGDGG 290

RESULT 19
T36689
Probable transpeptidase - Streptomyces coelicolor (fragment)
C;Species: Streptomyces coelicolor
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T36689
R;Murphy, L.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, April 1999
A;Reference number: Z21597
A;Accession: T36689
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-366 <MUR>
A;Cross-references: UNIPROT:Q8CJW2; UNIPARC:UPI000017AE05; EMBL:AL049731; PIDN:CAB41728.
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SCH66.01

Query Match 87.5%; Score 35; DB 2; Length 366;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
||| |||||
Db 339 GGTGDDG 345

RESULT 20
T04347
teosinte branched1 protein - maize (fragment)
C;Species: Zea mays (maize)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T04347
R;Doebley, J.; Stec, A.O.; Hubbard, L.
Nature 386, 485-488, 1997
A;Title: The evolution of apical dominance in maize.
A;Reference number: Z15300; MUID:97242406; PMID:9087405
A;Accession: T04347
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-382 <DOE>
A;Cross-references: UNIPROT:Q93WE3; UNIPROT:Q04170; UNIPARC:UPI000009FD45; EMBL:U94494;
A;Experimental source: strain Pioneer Inbred Ap9
C;Genetics:
A;Gene: tbt1
A;Map position: 1.09

Query Match 87.5%; Score 35; DB 2; Length 382;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
||| |||||
Db 353 GGYGDGG 359

RESULT 21
S01820

A;Experimental source: clone C47E8
C;Genetics:
A;Gene: CESP:C47E8.3
A;Map position: 5
A;Introns: 97/2; 194/3; 231/2

Query Match 87.5%; Score 35; DB 2; Length 317;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
||| |||||
Db 284 GGAGDGG 290

RESULT 19
T36689
Probable transpeptidase - Streptomyces coelicolor (fragment)
C;Species: Streptomyces coelicolor
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T36689
R;Murphy, L.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, April 1999
A;Reference number: Z21597
A;Accession: T36689
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-366 <MUR>
A;Cross-references: UNIPROT:Q8CJW2; UNIPARC:UPI000017AE05; EMBL:AL049731; PIDN:CAB41728.
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SCH66.01

Query Match 87.5%; Score 35; DB 2; Length 366;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
||| |||||
Db 339 GGTGDDG 345

RESULT 20
T04347
teosinte branched1 protein - maize (fragment)
C;Species: Zea mays (maize)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T04347
R;Doebley, J.; Stec, A.O.; Hubbard, L.
Nature 386, 485-488, 1997
A;Title: The evolution of apical dominance in maize.
A;Reference number: Z15300; MUID:97242406; PMID:9087405
A;Accession: T04347
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-382 <DOE>
A;Cross-references: UNIPROT:Q93WE3; UNIPROT:Q04170; UNIPARC:UPI000009FD45; EMBL:U94494;
A;Experimental source: strain Pioneer Inbred Ap9
C;Genetics:
A;Gene: tbt1
A;Map position: 1.09

Query Match 87.5%; Score 35; DB 2; Length 382;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
||| |||||
Db 353 GGYGDGG 359

RESULT 21
S01820

glycine-rich cell wall protein 1.8 precursor - kidney bean
C;Species: Phaseolus vulgaris (kidney bean)
C;Date: 30-Sep-1989 #sequence_revision 19-May-1994 #text_change 09-Jul-2004
C;Accession: S01820
R;Keller, B.; Sauer, N.; Lamb, C.J.
EMBO J. 7, 3625-3633, 1988
A;Title: Glycine-rich cell wall proteins in bean: Gene structure and association of the
A;Reference number: S01820; MUID:89091109; PMID:3208742
A;Accession: S01820
A;Molecule type: DNA
A;Residues: 1-465 <KEL>
A;Cross-references: UNIPROT:P10496; UNIPARC:UPI000012BB05; EMBL:X13596; NID:g21002; PID:
C;Comment: This protein is enriched in the cell wall fraction of young hypocotyls and o
tyls.
C;Superfamily: Much of the sequence consists of tandemly repeated 22-residue segments with
C;Superfamily: Phaseolus glycine-rich cell wall protein 1.8
C;Keywords: cell wall; structural protein; tandem repeat
F1-30/Domain: signal sequence #status predicted <SIG>
F;31-465/Product: glycine-rich cell wall protein 1.8 #status predicted <MAT>

Query Match 87.5%; Score 35; DB 1; Length 465;
Best Local Similarity 85.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
||| |||||
Db 418 GGYGDGG 424

RESULT 22
WMBEH2
UL32 protein - human herpesvirus 1 (strain 17)
C;Species: human herpesvirus 1
C;Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 09-Jul-2004
C;Accession: E30085
R;McGeoch, D.J.; Dalrymple, M.A.; Davison, A.J.; Dolan, A.; Frame, M.C.; McNab, D.; Per
J. Gen. Virol. 69, 1531-1574, 1988
A;Title: The complete DNA sequence of the long unique region in the genome of herpes si
A;Reference number: A30083; MUID:88274327; PMID:2839594
A;Accession: E30085
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-596 <MCG>
A;Cross-references: UNIPROT:P10216; UNIPARC:UPI0000137B91; GB:X14112; NID:g1944536; PID:
C;Genetics:
A;Gene: UL32
C;Superfamily: varicella-zoster virus gene 26 protein
C;Keywords: glycoprotein
F;245,451,452/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 87.5%; Score 35; DB 1; Length 596;
Best Local Similarity 85.7%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
||| |||||
Db 82 GGAGDGG 88

RESULT 23
D84675
hypothetical protein At2g27660 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: D84675
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon, L
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: D84675
A;Status: preliminary

A:Molecule type: DNA
A:Residues: 1-718 <STO>
A:Cross-references: UNIPROT:Q9ZUW8; UNIPARC:UPI00000A8A4D; GB:AE002093; NID:g3860266; PI
C:Genetics:
A:Gene: At2g27660
A:Map position: 2

Query Match 87.5%; Score 35; DB 2; Length 718;
Best Local Similarity 85.7%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGIGDGG 7
Db 701 GGAGDGG 707

RESULT 24
T52510
hypothetical protein B2J23.60 [imported] - Neurospora crassa
C:Species: Neurospora crassa
C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000
C:Accession: T52510
R:Schulte, U.; Aign, V.; Hohnel, J.; Brandt, P.; Partmann, B.; Holland, R.; Nyakatura, R.
submitted to the Protein Sequence Database, September 2000
A:Reference number: Z26053
A:Accession: T52510
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-719 <SCH>
A:Cross-references: UNIPARC:UPI000017B49D; EMBL:AL442164; GSPDB:GN00116; NCSP:B2J23.60
A:Experimental source: BAC clone B2J23; strain OR74A
C:Genetics:
A:Gene: NCSP:B2J23.60
A:Map position: 6
A:Introns: 349/1; 601/1

Query Match 87.5%; Score 35; DB 2; Length 719;
Best Local Similarity 85.7%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGIGDGG 7
Db 336 GGAGDGG 342

RESULT 25
B70807
hypothetical glycine-rich protein Rv3512 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: B70807
R:Colle, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: B70807
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1079 <COL>
A:Cross-references: UNIPARC:UPI00000D3AEA; GB:AL022022; GB:AL123456; NID:g3261554; PIDN:
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: Rv3512
C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;

Query Match 87.5%; Score 35; DB 2; Length 1079;
Best Local Similarity 85.7%; Pred. No. 3.8e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGIGDGG 7

Db 1048 GGAGDGG 1054

RESULT 26
P95952

probable membrane-anchored protein [imported] - Sinorhizobium meliloti (strain 1021) m
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C:Accession: P95952
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hern
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A:Title: The complete sequence of the 1,683-Kb pSymB megaplasmid from the N2-fixing en
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: P95952
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-234 <KUR>
A:Cross-references: UNIPROT:P96433; UNIPARC:UPI00000CA9B5; GB:AL591985; PIDN:CAC49286.
A:Experimental source: strain 1021, megaplasmid pSymB
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Huble
Pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaui
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh,
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: expB5; SMB211310
A:Genome: plasmid

Query Match 85.0%; Score 34; DB 2; Length 234;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GIGDGG 7
Db 72 GIGDGG 77

RESULT 27

G71208
hypothetical protein PH1935 - Pyrococcus horikoshii

C:Species: Pyrococcus horikoshii
C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 12-Jul-2004
C:Accession: G71208
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Seki
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguci
DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic
A:Reference number: A71000; MUID:98344137; PMID:9679194
A:Accession: G71208
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-266 <KAW>
A:Cross-references: UNIPROT:O59598; UNIPARC:UPI0000063186; GB:AP000007; NID:g3236134;
A:Experimental source: strain OT3
A:Note: this accession replaces an interim accession for a sequence replaced by GenBank
C:Genetics:
A:Gene: PH1935

Query Match 85.0%; Score 34; DB 2; Length 266;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GIGDGG 7
Db 159 GIGDGG 164

RESULT 28

H75021
hypothetical protein PAB1225 - Pyrococcus abyssi (strain Orsay)
C:Species: Pyrococcus abyssi
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C:Accession: H75021
R:anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome structure
A:Reference number: A75001
A:Accession: H75021
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-268 <KAW>
A:Cross-references: UNIPROT:Q9UXZ8; UNIPARC:UPI00000634C2; GB:AJ248288; GB:AL096836; NID:
A:Experimental source: strain Orsay
C:Genetics:
A:Gene: PAB1225
C:Superfamily: Pyrococcus horikoshii hypothetical protein PH1935
Query Match 85.0%; Score 34; DB 2; Length 268;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 GGIGDG 7
DB 161 GGIGDG 166
RESULT 29
H97338
dioxxygenase related to 2-nitropropane dioxxygenase [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C:Accession: H97338
R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: H97338
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-310 <KUR>
A:Cross-references: UNIPROT:Q97DM4; UNIPARC:UPI000000CA899; GB:AE001437; PIDN:AAK81499.1;
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC3576
Query Match 85.0%; Score 34; DB 2; Length 310;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGIGDG 6
DB 170 GGIGDG 175
RESULT 30
G83979
hypothetical protein BH2639 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C:Accession: G83979
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: G83979
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-325 <SPO>
A:Cross-references: UNIPROT:Q9K9K6; UNIPARC:UPI000000C3F73; GB:AF001516; GB:BA000004; NID:
A:Experimental source: strain C-125

C:Genetics:
A:Gene: BH2639
Query Match 85.0%; Score 34; DB 2; Length 325;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGIGDG 6
DB 175 GGIGDG 180
RESULT 31
CEECFX
cell division protein ftsX - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Jul-2004
C:Accession: S03132; S47681; A65143
R:Gill, D.R.; Hatfull, G.F.; Salmond, G.P.C.
Mol. Gen. Genet. 205, 134-145, 1986
A:Title: A new cell division operon in Escherichia coli.
A:Reference number: S03129; MUID:87089083; PMID:3025556
A:Accession: S03132
A:Molecule type: DNA
A:Residues: 1-352 <GIL>
A:Cross-references: UNIPROT:P10122; UNIPARC:UPI000012ACFE; EMBL:X04398; NID:g41496; PIDN:
R:Plunkett, G.
submitted to the EMBL Data Library, March 1994
A:Reference number: S47666
A:Accession: S47681
A:Molecule type: DNA
A:Residues: 1-352 <PLU>
A:Cross-references: UNIPARC:UPI000012ACFE; EMBL:U00039; NID:g466582; PIDN:AAH18437.1; P
R:Blattner, P.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: A65143
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-352 <BLAT>
A:Cross-references: UNIPARC:UPI000012ACFE; GB:AE000422; GB:U00096; NID:g1789868; PIDN:AJ
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: ftsX
A:Map position: 76 min
C:Superfamily: cell division protein ftsX
C:Keywords: cell division
Query Match 85.0%; Score 34; DB 1; Length 352;
Best Local Similarity 85.7%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGIGDG 7
DB 25 GGSGDGG 31
RESULT 32
T46716
hypothetical protein L4326.05 [imported] - Leishmania major
C:Species: Leishmania major
C:Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 09-Jul-2004
C:Accession: T46716
R:Volckaert, G.; Ivens, A.C.; Lawson, D.; Quail, M.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, December 1999
A:Reference number: Z23137
A:Accession: T46716
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-415 <VOL>
A:Cross-references: UNIPROT:Q9U153; UNIPARC:UPI000007CE11; EMBL:AL121861; PIDN:CAB58381

A:Experimental source: strain Friedlin
C:Genetics:
A:Note: L4326.05

Query Match 85.0%; Score 34; DB 2; Length 415;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GIGDGG 7
DB 226 GIGDGG 231

RESULT 33
T12583
oleoyl-[acyl-carrier-protein] hydrolase (EC 3.1.2.14) - common sunflower
C:Species: Helianthus annuus (common sunflower)
C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 09-Jul-2004
C:Accession: T12583
R:Slabaugh, M.B.; Katengam, S.; Hongtrakul, V.; Knapp, S.J.
submitted to the EMBL Data Library, December 1997
A:Description: Sunflower FatB thioesterase cDNA.
A:Reference number: Z17543
A:Accession: T12583
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-430 <SLA>
A:Cross-references: UNIPROT:O48568; UNIPARC:UPI00000A01AD; EMBL:AF036565; NID:g2687424;
A:Experimental source: cultivar Mammoth
C:Genetics:
A:Gene: FatB1
C:Superfamily: Arabidopsis oleoyl-[acyl-carrier-protein] hydrolase
C:Keywords: thiolester hydrolase

Query Match 85.0%; Score 34; DB 2; Length 430;
Best Local Similarity 85.7%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
DB 376 GGIGDSG 382

RESULT 34
T46437
hypothetical protein DKFZp434K0126.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C:Accession: T46437
R:Ansorge, W.; Wirkner, U.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A:Reference number: Z23028
A:Accession: T46437
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-431 <AAA>
A:Cross-references: UNIPROT:Q9NTC4; UNIPARC:UPI000006D912; EMBL:AL137377
A:Experimental source: adult testis; clone DKFZp434K0126
C:Genetics:
A:Note: DKFZp434K0126.1
C:Superfamily: transitional endoplasmic reticulum ATPase; FtsH/SEC18/CDC48-type ATP-bind

Query Match 85.0%; Score 34; DB 2; Length 431;
Best Local Similarity 85.7%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
DB 213 GNIGDGG 219

RESULT 35
AC2306

hypothetical protein all4002 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AC2306
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigun
Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata,
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium A
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AC2306
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-476 <KUR>
A:Cross-references: UNIPROT:Q8YQ34; UNIPARC:UPI00000CE9A3; GB:BA000019; PIDN:BA075701
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all4002
C:Superfamily: photosystem II chlorophyll a-binding protein pabc

Query Match 85.0%; Score 34; DB 2; Length 476;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GIGDGG 7
DB 76 GIGDGG 81

RESULT 36
T17478
hypothetical protein PCZA361.11 - Amycolatopsis orientalis
C:Species: Amycolatopsis orientalis
C:Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 09-Jul-2004
C:Accession: T17478
R:Van Wageningen, A.; Kirkpatrick, P.; Williams, D.; Harris, B.; Kershaw, J.; Lennard,
Chem. Biol. 3, 155-162, 1998
A:Title: Sequencing and analysis of genes involved in the biosynthesis of a vancomycin
A:Reference number: Z18804
A:Accession: T17478
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-495 <VAN>
A:Cross-references: UNIPROT:O52799; UNIPARC:UPI00000B7C0A; EMBL:AJ223998; NID:e1251208,

Query Match 85.0%; Score 34; DB 2; Length 495;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GIGDGG 7
DB 173 GIGDGG 178

RESULT 37
H70839
hypothetical glycine-rich protein Rv0109 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: H70839
R:Conor, R.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon,
Rajandream, M.A.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genomic
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: H70839
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-496 <COL>
A:Cross-references: UNIPARC:UPI00000D3AEE; GB:AL021926; GB:AL123456; NID:g3261520; PIDN
A:Experimental source: strain H37RV

C;Genetics:
A;Gene: Rv0109
C;Superfamily: Phaseolus glycine-rich cell wall protein 1.8

Query Match 85.0%; Score 34; DB 2; Length 496;
Best Local Similarity 71.4%; Pred. No. 2.6e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
|||:|:
Db 243 GGVNGG 249

RESULT 38

T33024
hypothetical protein K07H8.2 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T33024
R;Fulton, B.; Hawkins, J.; Gattung, S.; Wohldmann, P.; Elliott, G.
submitted to the EMBL Data Library, February 1998
A;Description: The sequence of C. elegans cosmid K07H8.
A;Reference number: Z21264
A;Accession: T33024
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-507 <FUL>
A;Cross-references: UNIPROT:Q45182; UNIPARC:UPI000080E7E; EMBL:AF047659; PIDN:AAC04427.
A;Experimental source: strain Bristol N2; clone K07H8
C;Genetics:
A;Gene: CBSP:K07H8.2
A;Map position: 4
A;Introns: 20/3; 55/2; 93/3; 152/3; 447/3

Query Match 85.0%; Score 34; DB 2; Length 507;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDG 6
|||||
Db 5 GGIGDG 10

RESULT 39

B49131
ryanodine receptor Dry / intracellular Ca(2+)-release channel homolog - fruit fly (Dros

C;Species: Drosophila melanogaster
C;Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 11-Jun-1999
C;Accession: B49131
R;Hasan, G.; Rosbash, M.
Development 116, 967-975, 1992
A;Title: Drosophila homologs of two mammalian intracellular Ca(2+)-release channels: ide
A;Reference number: A49131; MUID:93202018; PMID:1338312
A;Accession: B49131
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-508 <HAS>
A;Cross-references: UNIPARC:UPI000017794A
A;Experimental source: embryo
A;Note: sequence extracted from NCBI backbone (NCBIP:127825)
C;Genetics:

A;Gene: FlyBase:Rya-r76CD
C;Cross-references: FlyBase:FBgn0010268
C;Superfamily: ryanodine receptor; transcription initiation factor sigma region 1 homolog
Query Match 85.0%; Score 34; DB 2; Length 508;
Best Local Similarity 85.7%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
|||||
Db 83 GGSDDG 89

RESULT 40

T40856

probable nucleotide phosphatase - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T40856
R;Rampberger, U.; Pohl, T.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, October 1999

A;Reference number: Z21952
A;Accession: T40856
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-572 <RAM>

A;Cross-references: UNIPROT:Q9USP2; UNIPARC:UPI000006B4B4; EMBL:AL121783; PIDN:CAB57847
A;Experimental source: strain 972h-; cosmid cille10
C;Genetics:
A;Gene: SPDB:SPCC11E10.05c
A;Map position: 3

Query Match 85.0%; Score 34; DB 2; Length 572;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GIGDGG 7
|||||
Db 44 GIGDGG 49

RESULT 41

T48224

probable homeodomain protein - Arabidopsis thaliana

N;Alternate names: protein T7H20.80
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C;Accession: T48224
R;Bevan, M.; Peters, S.A.; van Staveren, M.; Dirkse, W.; Stiekema, W.; Bancroft, I.; Me
submitted to the Protein Sequence Database, March 2000
A;Reference number: Z24488
A;Accession: T48224
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-575 <BEV>
A;Cross-references: UNIPROT:Q9LZM8; UNIPARC:UPI00000A5BD9; EMBL:AL162508
A;Experimental source: cultivar Columbia, BAC clone T7H20
C;Genetics:
A;Map position: 5
A;Introns: 240/3; 374/2; 394/3
A;Note: T7H20.80

Query Match 85.0%; Score 34; DB 2; Length 575;
Best Local Similarity 85.7%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
|||||
Db 541 GSGDGG 547

RESULT 42

S23391

transcription factor E12 - African clawed frog

N;Alternate names: DNA-binding protein E12
C;Species: Xenopus laevis (African clawed frog)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: S23391
R;Rashbass, J.; Taylor, M.V.; Gurdon, J.B.
EMBO J. 11, 2361-2390, 1992
A;Title: The DNA-binding protein E12 co-operates with XMyoD in the activation of muscle

A;Reference number: S23391; MUID:92347333; PMID:1322293
A;Accession: S23391
A;Molecule type: mRNA
A;Residues: 1-658 <RAS>

A:Cross-references: UNIPROT:Q01978; UNIPARC:UPI000000BABB; EMBL:X66959; NID:G65223; PIDN: C:Superfamily: human transcription factor 3
C:Keywords: transcription factor

Query Match 85.0%; Score 34; DB 1; Length 658;
Best Local Similarity 85.7%; Pred. No. 3.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
DB 233 GGIGGG 239

RESULT 43
T14605
probable cell division cycle protein - Trypanosoma cruzi
C:Species: Trypanosoma cruzi
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T14605
R:Anderson, B.; Aelund, L.; Pettersson, U.
A:Submitted to the EMBL Data Library, March 1998
A:Description: 93.4 kb of complete sequence from chromosome 3 of Trypanosoma cruzi.
A:Reference number: Z18159
A:Accession: T14605
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-723 <AND>
A:Cross-references: UNIPROT:Q61098; UNIPARC:UPI000008374B; EMBL:AF052832; NID:G3063540;
C:Genetics:
A:Map position: 3

Query Match 85.0%; Score 34; DB 2; Length 723;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GIGDGG 7
DB 689 GIGDGG 694

RESULT 44
C70974
hypochemical glycine-rich protein Rv3388 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 07-Mar-2003
C:Accession: C70974
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: C70974
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-731 <COL>
A:Cross-references: UNIPARC:UPI000000D3ADA; GB:AL123456; NID:G3242262; PIDN:
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: Rv3388
C:Superfamily: elastin related uncharacterized glycine-rich protein, PE motif containing

Query Match 85.0%; Score 34; DB 2; Length 731;
Best Local Similarity 71.4%; Pred. No. 3.8e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
DB 438 GGVGNGG 444

RESULT 45

S19738
transitional endoplasmic reticulum ATPase (EC 3.6.1.-) 97K chain - African clawed frog
N:Alternate names: CDC48 homolog; valosin-containing protein
N:Contains: adeninetriphosphatase (EC 3.6.1.3) [validated]
C:Species: Xenopus laevis (African clawed frog)
C:Date: 07-May-1993 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: S19738; S12537
R:Peterson, J.
A:Submitted to the EMBL Data Library, July 1991
A:Reference number: S19738
A:Accession: S19738
A:Molecule type: mRNA
A:Residues: 1-805 <PET>
A:Cross-references: UNIPROT:P23787; UNIPARC:UPI0000136BCA; EMBL:X54240; NID:G64965; PI
R:Peterson, J.M.; Walsh, M.J.; Franke, W.W.
EMBO J. 9, 1757-1767, 1990
A:Title: An abundant and ubiquitous homo-oligomeric ring-shaped ATPase particle related
A:Reference number: S12537; MUID:90269209; PMID:2140770
A:Accession: S12537
A:Molecule type: mRNA
A:Residues: 442-805 <PET2>
A:Cross-references: UNIPARC:UPI0000172C3A; EMBL:X54240
A:Note: the authors translated the codon GGG for residue 455 as Ala
C:Complex: homooligomer [validated, MUID:90269209]
C:Function:
A:Description: may be the transitional endoplasmic reticulum ATPase
C:Superfamily: transitional endoplasmic reticulum ATPase; FtsH/SEC18/CDC48-type ATP-bl.
C:Keywords: ATP; duplication; hydrolase; nucleotide binding; P-loop
F:218-425/Domain: FtsH/SEC18/CDC48-type ATP-binding domain homology <VAT1>
F:245-252/Region: nucleotide-binding motif A (P-loop)
F:491-701/Domain: FtsH/SEC18/CDC48-type ATP-binding domain homology <VAT2>
F:518-525/Region: nucleotide-binding motif A (P-loop)

Query Match 85.0%; Score 34; DB 1; Length 805;
Best Local Similarity 85.7%; Pred. No. 4.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
DB 588 GNIGDGG 594

RESULT 46
VPPG
transitional endoplasmic reticulum ATPase - pig
N:Alternate names: CDC48 homolog; TER ATPase; valosin-containing protein
N:Contains: ATPase (EC 3.6.1.-)
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 13-Aug-1986 #sequence_revision 16-Feb-1996 #text_change 19-Jan-2001
C:Accession: A26360; A01627
R:Koller, K.J.; Brownstein, M.J.
Nature 325, 542-545, 1987
A:Title: Use of a cDNA clone to identify a supposed precursor protein containing valos
A:Reference number: A26360; MUID:87115844; PMID:3468358
A:Accession: A26360
A:Molecule type: mRNA
A:Residues: 1-806 <COL>
A:Cross-references: UNIPARC:UPI0000172C35; GB:M30143
R:Schmidt, W.E.; Mutt, V.; Carlquist, M.; Kratzin, H.; Conlon, J.M.; Creutzfeldt, W.
FEBS Lett. 191, 264-268, 1985
A:Title: Valosin: isolation and characterization of a novel peptide from porcine intest
A:Reference number: A01627; MUID:86030678; PMID:4054310
A:Accession: A01627
A:Molecule type: protein
A:Residues: 493-517 <SCH>
A:Cross-references: UNIPARC:UPI0000172C36
C:Comment: The peptide valosin is probably a breakdown product with no physiological s
C:Superfamily: transitional endoplasmic reticulum ATPase; FtsH/SEC18/CDC48-type ATP-bl.
C:Keywords: ATP; duplication; hydrolase; nucleotide binding; P-loop; phosphoprotein
F:218-425/Domain: FtsH/SEC18/CDC48-type ATP-binding domain homology <VAT1>
F:245-252/Region: nucleotide-binding motif A (P-loop)
F:491-701/Domain: FtsH/SEC18/CDC48-type ATP-binding domain homology <VAT2>
F:518-525/Region: nucleotide-binding motif A (P-loop)

Query Match 85.0%; Score 34; DB 1; Length 806;
Best Local Similarity 85.7%; Pred. No. 4.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
| | | | |
Db 588 GNIGDGG 594

RESULT 47
A55190
transitional endoplasmic reticulum ATPase (EC 3.6.1.1-) [validated] - rat
N;Alternate names: CDC48 homolog; TER ATPase; valosin-containing protein
C;Species: Rattus norvegicus (Norway rat)
C;Date: 27-Jan-1995 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C;Accession: A55190
R;Zhang, L.; Ashendel, C.L.; Becker, G.W.; Morre, D.J.
J. Cell Biol. 127, 1871-1883, 1994
A;Title: Isolation and characterization of the principal ATPase associated with transiti
A;Reference number: A55190; MUID:95105231; PMID:7806566
A;Accession: A55190
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-806 <ZHA>
A;Cross-references: UNIPROT:P46462; UNIPARC:UPI0000136BC9; GB:U11760; NID:G641972; PIDN:
A;Experimental source: liver
C;Complex: hexamer of six 100K chains [validated, MUID:95105231]; ring-shaped structure
C;Function:
A;Description: EC 3.6.1.1-; ATPase [validated, MUID:95105231]
A;Note: the enzyme is phosphorylated in the presence of magnesium
C;Superfamily: transitional endoplasmic reticulum ATPase; FtsH/SEC18/CDC48-type ATP-bind
C;Keywords: ATP; duplication; hydrolase; nucleotide binding; P-loop; phosphoprotein
F;218-425/Domain: FtsH/SEC18/CDC48-type ATP-binding domain homology <VAT1>
F;245-252/Region: nucleotide-binding motif A (P-loop)
F;491-701/Domain: FtsH/SEC18/CDC48-type ATP-binding domain homology <VAT2>
F;518-525/Region: nucleotide-binding motif A (P-loop)

Query Match 85.0%; Score 34; DB 1; Length 806;
Best Local Similarity 85.7%; Pred. No. 4.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
| | | | |
Db 588 GNIGDGG 594

RESULT 48
S25197
transitional endoplasmic reticulum ATPase - mouse
N;Alternate names: CDC48 homolog; TER ATPase; valosin-containing protein
N;Contains: ATPase (EC 3.6.1.1-)
C;Species: Mus musculus (house mouse)
C;Date: 07-May-1993 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C;Accession: S25197; S30329
R;Egerton, M.; Ahe, O.R.; Chen, D.; Druker, B.J.; Burgess, W.H.; Samelson, L.E.
EMBO J. 11, 3533-3540, 1992
A;Title: VCP, the mammalian homolog of cdc48, is tyrosine phosphorylated in response to
A;Reference number: S25197; MUID:93010943; PMID:1382975
A;Accession: S25197
A;Molecule type: mRNA
A;Residues: 1-806 <EGE>
A;Cross-references: UNIPROT:Q01853; UNIPARC:UPI0000029436; EMBL:Z14044; NID:G55216; PIDN:
A;Accession: S30329
A;Molecule type: protein
A;Residues: 20-40;295-309;425-438 <EG2>
A;Cross-references: UNIPARC:UPI0000172C37; UNIPARC:UPI0000172C38; UNIPARC:UPI0000172C39
C;Superfamily: transitional endoplasmic reticulum ATPase; FtsH/SEC18/CDC48-type ATP-bind
C;Keywords: ATP; duplication; hydrolase; nucleotide binding; P-loop; phosphoprotein
F;218-425/Domain: FtsH/SEC18/CDC48-type ATP-binding domain homology <VAT1>
F;245-252/Region: nucleotide-binding motif A (P-loop)
F;491-701/Domain: FtsH/SEC18/CDC48-type ATP-binding domain homology <VAT2>
F;518-525/Region: nucleotide-binding motif A (P-loop)

Query Match 85.0%; Score 34; DB 1; Length 806;
Best Local Similarity 85.7%; Pred. No. 4.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
| | | | |
Db 588 GNIGDGG 594

RESULT 49
T02243
probable transitional endoplasmic reticulum ATPase - human
C;Species: Homo sapiens (man)
C;Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 09-Jul-2004
C;Accession: T02243
R;Lamerdin, J.E.; McCready, P.M.; Skowronski, E.; Adamson, A.W.; Burkhardt-Schultz, K.; (; Christensen, M.; Georgescu, A.; Avila, J.; Liu, S.; Attix, C.; Andreise, T.; Trankhe:
submitted to the EMBL Data Library, March 1998
A;Authors: Duarte, S.; Lucas, S.; Bruce, R.; Thomas, P.; Quan, G.; Kronmiller, B.; Arel:
A;Description: Sequence analysis of a human P1 clone containing the XRCC9 DNA repair ge
A;Reference number: Z14637
A;Accession: T02243
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-806 <LAM>
A;Cross-references: UNIPROT:P55072; UNIPARC:UPI000005FB2E; EMBL:AC004472; NID:G2984582;
C;Genetics:
A;Map position: 9
A;Intons: 6/2; 43/3; 101/2; 149/1; 192/3; 236/3; 271/1; 315/3; 361/1; 398/3; 453/3; 49:
C;Superfamily: transitional endoplasmic reticulum ATPase; FtsH/SEC18/CDC48-type ATP-bin
F;491-701/Domain: FtsH/SEC18/CDC48-type ATP-binding domain homology <VAT>

Query Match 85.0%; Score 34; DB 2; Length 806;
Best Local Similarity 85.7%; Pred. No. 4.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
| | | | |
Db 588 GNIGDGG 594

RESULT 50
S70012
methionine-S-oxide reductase (EC 1.8.4.5) precursor - Rhodobacter sphaeroides
N;Alternate names: dimethyl sulfoxide reductase
C;Species: Rhodobacter sphaeroides
C;Date: 06-Dec-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C;Accession: S70012; JC4322; PC4102; S66255; S66263
R;Hilton, J.C.; Rajagopalan, K.V.
Biochim. Biophys. Acta 1294, 111-114, 1996
A;Title: Molecular cloning of dimethyl sulfoxide reductase from Rhodobacter sphaeroides
A;Reference number: S70012; MUID:96248388; PMID:8645727
A;Accession: S70012
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-822 <HIL>
A;Cross-references: UNIPROT:Q57366; UNIPARC:UPI0000129504; EMBL:L46851; NID:gl1161235; P:
A;Experimental source: strain IL106
R;Yamamoto, I.; Wada, N.; Ujiiye, T.; Tachibana, M.; Matsuzaki, M.; Kajiwara, H.; Watan
Biosci. Biotechnol. Biochem. 59, 1850-1855, 1995
A;Title: Cloning and nucleotide sequence of the gene encoding dimethyl sulfoxide reduct
A;Reference number: JC4322; MUID:96068928; PMID:8534974
A;Accession: JC4322
A;Molecule type: DNA
A;Residues: 1-822 <YAM>
A;Cross-references: UNIPARC:UPI0000129504; DDBJ:D38634; NID:gl1177801; PIDN:BA007615.1;
A;Molecule type: protein
A;Accession: PC4102
A;Residues: 43-53;269-284;286-293;427-440;617-621;623-626;'ED';644-656 <YA2>
A;Cross-references: UNIPARC:UPI0000178AB7; UNIPARC:UPI0000178AB8; UNIPARC:UPI0000178AB9
ABE
A;Experimental source: IL106

R:Barber, M.J.; van Valkenburgh, H.; Trimboli, A.J.; Pollock, V.V.; Neame, P.J.; Bastian Arch. Biochem. Biophys. 320, 266-275, 1995
A:Title: The amino acid sequence of Rhodobacter sphaeroides dimethyl sulfoxide reductase
A:Reference number: S66255; MUID:95351787; PMID:7625833
A:Accession: S66255
A>Status: preliminary
A:Molecule type: protein
A:Residues: 43-87;806-813,'F',815-816,'P',818 <BAR>
A:Cross-references: UNIPARC:UPI0000178ABF; UNIPARC:UPI0000178AC0
A:Accession: S66263
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 53-158,'R',160-184,'V',186-191,'G',193-198,'Y',200-370,'MAQ',374-421,'A',423
A:Cross-references: UNIPARC:UPI0000178ABF; UNIPARC:UPI0000178AC0; EMBL:025037
A:Experimental source: strain IL 106
A>Note: the authors translated the codon NAT for residue 129 as Lys, GTC for residue 425
C:Genetics:
A:Gene: dmsA; dsaA
C:Function:
A:Description: catalyzes the reduction of many oxide compounds
A>Note: cofactor molybdenum; is synthesized inducibly in the presence of dimethyl sulfox
C:Superfamily: trimethylamine-N-oxide reductase
C:Keywords: molybdenum; monomer; oxidoreductase; periplasmic space
F:1-42/Domain: signal sequence #status predicted <SIG>
F:43-822/Product: methionine-S-oxide reductase #status predicted <MAT>

Query Match 85.0%; Score 34; DB 2; Length 822;
Best Local Similarity 85.7%; Pred. No. 4.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
DB 416 GGISDGG 422

Search completed: February 27, 2006, 09:50:37
Job time : 43 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: February 27, 2006, 09:42:49 ; Search time 234 Seconds
(without alignments)
21.106 Million cell updates/sec

Title: US-10-027-015B-2

Perfect score: 40

Sequence: 1 GGIGGG 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : UniProt 05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	40	100.0	195	2	Q4FXV3_LEIMA	Q4fxv3 leishmania
2	40	100.0	256	2	Q9RSI7_DEIRA	Q9rsi7 deinococcus
3	40	100.0	289	2	Q7BTU0_YERPE	Q7btu0 yersinia pe
4	40	100.0	314	2	Q935F2_SALTI	Q935f2 salmonella
5	40	100.0	314	2	Q9RIC9_YERPE	Q9ric9 yersinia pe
6	40	100.0	331	2	Q68765_YERPE	Q68765 yersinia pe
7	40	100.0	331	2	Q65AK6_YERPE	Q65ak6 yersinia pe
8	40	100.0	351	2	Q43494_LUPAN	Q43494 lupinus ang
9	40	100.0	351	2	Q43878_LUPAN	Q43878 lupinus ang
10	40	100.0	395	2	Q6AL55_DESPS	Q6al55 desulfotale
11	40	100.0	425	2	Q24779_PSEAE	Q24779 pseudomonas
12	40	100.0	425	2	Q91083_PSEAE	Q91083 pseudomonas
13	40	100.0	426	2	Q5Q559_PSEFL	Q5q559 pseudomonas
14	40	100.0	527	1	DRTS_TRYBB	Q27783 trypanosoma
15	40	100.0	527	2	Q582G3_9TRYP	Q582g3 trypanosoma
16	40	100.0	742	1	DNML_MOUSE	Q8k1m6 mus musculus
17	40	100.0	755	1	DNML_RAT	Q35303 rattus norv
18	40	100.0	984	2	Q4NTR7_DELT	Q4ntr7 anaeromyxob
19	40	100.0	1090	1	NIT4_NEUCR	P28349 neurospora
20	39	97.5	162	2	Q5A4Z5_CANAL	Q5a4z5 candida alb
21	39	97.5	261	2	Q4TID1_TETNG	Q4tid1 tetraodon n
22	39	97.5	294	2	Q6OAS9_METCA	Q6oas9 methylococc
23	39	97.5	333	2	Q7XHX5_ORYSA	Q7xhx5 oryza sativ
24	39	97.5	334	2	Q61EN4_ORYSA	Q61en4 oryza sativ
25	39	97.5	439	2	Q88R66_PSEPK	Q88r66 pseudomonas
26	39	97.5	525	2	Q9NXG7_HUMAN	Q9nxg7 homo sapien
27	39	97.5	1291	2	Q8LNH8_ORYSA	Q8lnh8 oryza sativ
28	39	97.5	1356	2	Q7SBF8_NEUCR	Q7sbf8 neurospora
29	38	95.0	214	2	Q9LG66_ORYSA	Q9lg66 oryza sativ
30	38	95.0	228	2	Q5BIAS_DROME	Q5bias drosophila
31	38	95.0	273	2	Q75I35_ORYSA	Q75i35 oryza sativ

ALIGNMENTS

32	38	95.0	300	2	Q6YUM4_ORYSA	Q6yum4 oryza sativ
33	38	95.0	427	2	Q94534_DROME	Q94534 drosophila
34	38	95.0	664	2	Q4H37_9DEIO	Q4h37 deinococcus
35	37	92.5	200	2	Q4H3G4_CIOIN	Q4h3g4 ciona intes
36	37	92.5	325	2	Q4H3G5_CIOIN	Q4h3g5 ciona intes
37	37	92.5	405	2	Q51P11_MAGGR	Q51p11 magnaporthe
38	37	92.5	417	2	Q7UXR4_RHOBA	Q7uxr4 rhodospirell
39	36	90.0	100	2	Q9DUC2_9VIRU	Q9duc2 torque teno
40	36	90.0	130	2	Q52747_ORYSA	Q52747 oryza sativ
41	36	90.0	223	2	Q9C6Y9_ARATH	Q9c6y9 arabidopsis
42	36	90.0	224	2	Q9A971_CAUCR	Q9a971 caulobacter
43	36	90.0	303	2	Q4P315_USTMA	Q4p315 ustilago ma
44	36	90.0	368	2	Q68842_STRAT	Q68842 streptomyce
45	36	90.0	368	2	Q91L6B9_STRAT	Q91l6b9 streptomyce
46	36	90.0	383	2	Q61561_ORYSA	Q61561 oryza sativ
47	36	90.0	405	2	Q69SQ7_ORYSA	Q69sq7 oryza sativ
48	36	90.0	408	2	Q5RJ37_BRARE	Q5rj37 brachydanio
49	36	90.0	424	2	Q4ZLL5_PSESY	Q4zll5 pseudomonas
50	36	90.0	424	2	Q87UB4_PSESM	Q87ub4 pseudomonas
51	36	90.0	456	2	Q82JF7_STRAW	Q82jf7 streptomyce
52	36	90.0	524	2	Q9IED3_9HIV1	Q9ied3 human immun
53	36	90.0	526	1	K1C10_BOVIN	P06394 bos taurus
54	36	90.0	561	2	Q4RRB2_TETNG	Q4rrb2 tetraodon n
55	36	90.0	628	2	Q92213_ARCFU	Q92213 archaeoglob
56	36	90.0	690	2	Q55YH2_CRYNE	Q55yh2 cryptococcu
57	36	90.0	690	2	Q5KLG8_CRYNE	Q5klg8 cryptococcu
58	36	90.0	835	2	Q8LMJ1_ORYSA	Q8lmj1 oryza sativ
59	36	90.0	1278	2	Q82JRO_STRAW	Q82jro streptomyce
60	36	90.0	1891	2	Q97017_9CNID	Q97017 stylophora
61	36	90.0	4317	2	Q623Y1_CAEBR	Q623y1 caenorhabdi
62	36	90.0	4368	2	Q61851_CABEL	Q61851 caenorhabdi
63	36	90.0	11096	2	Q914W3_9ACTO	Q914w3 streptomyce
64	35	87.5	75	2	Q6UUK5_ORYSA	Q6uuk5 oryza sativ
65	35	87.5	83	2	Q5N815_ORYSA	Q5n815 oryza sativ
66	35	87.5	88	2	Q5Z4A5_ORYSA	Q5z4a5 oryza sativ
67	35	87.5	90	2	Q69X80_ORYSA	Q69x80 oryza sativ
68	35	87.5	91	2	Q69LL6_ORYSA	Q69ll6 oryza sativ
69	35	87.5	116	2	Q8XT51_RALSO	Q8xt51 raietonia s
70	35	87.5	117	2	Q94257_CABEL	Q94257 caenorhabdi
71	35	87.5	130	2	Q9AR21_ZRAMP	Q9ar21 zea mays (s
72	35	87.5	130	2	Q9AT82_MAIZE	Q9at82 zea mays (m
73	35	87.5	130	2	Q9AT84_MAIZE	Q9at84 zea mays (m
74	35	87.5	130	2	Q9AT87_ZEAMP	Q9at87 zea mays (s
75	35	87.5	130	2	Q9AT89_ZEAMP	Q9at89 zea mays (s
76	35	87.5	130	2	Q9AT91_ZEADI	Q9at91 zea diplope
77	35	87.5	132	2	Q9AR16_MAIZE	Q9ar16 zea mays (m
78	35	87.5	132	2	Q9AR17_ZEAMP	Q9ar17 zea mays (s
79	35	87.5	132	2	Q9AT83_MAIZE	Q9at83 zea mays (m
80	35	87.5	132	2	Q9AT85_MAIZE	Q9at85 zea mays (m
81	35	87.5	132	2	Q9AT86_ZEAMP	Q9at86 zea mays (s
82	35	87.5	132	2	Q9AT88_ZEAMP	Q9at88 zea mays (s
83	35	87.5	132	2	Q9AT90_ZEAMP	Q9at90 zea mays (s
84	35	87.5	132	2	Q631F1_BURPS	Q631f1 burkholderi
85	35	87.5	132	2	Q4V2V2_BURMA	Q4v2v2 burkholderi
86	35	87.5	154	2	Q77JT4_9VIRU	Q77jt4 ttv-like vi
87	35	87.5	154	2	Q99A06_9VIRU	Q99aq6 torque teno
88	35	87.5	167	2	Q8H8V4_ORYSA	Q8h8v4 oryza sativ
89	35	87.5	173	2	Q5QMP5_ORYSA	Q5qmp5 oryza sativ
90	35	87.5	185	2	Q5SIX3_ARATH	Q5six3 arabidopsis
91	35	87.5	193	2	Q93WZ6_PUPE	Q93wz6 prunus pers
92	35	87.5	194	2	Q96853_SCHHA	Q96853 schistosoma
93	35	87.5	199	2	Q651E7_ORYSA	Q651e7 oryza sativ
94	35	87.5	199	2	Q6K6W2_ORYSA	Q6k6w2 oryza sativ
95	35	87.5	200	2	Q50000_PUAR	Q50000 prunus arme
96	35	87.5	201	2	Q62160_ORYSA	Q62160 oryza sativ
97	35	87.5	204	2	Q7FA14_ORYSA	Q7fa14 oryza sativ
98	35	87.5	204	2	Q7X121_ORYSA	Q7x121 oryza sativ
99	35	87.5	228	2	Q8S0R0_ORYSA	Q8s0r0 oryza sativ
100	35	87.5	237	2	Q6AVL7_ORYSA	Q6avl7 oryza sativ

```
RESULT 1
Q4FXV3 LEIMA
ID Q4FXV3 LEIMA PRELIMINARY; PRT; 195 AA.
AC Q4FXV3;
DT 13-SEP-2005 (TReMBLrel. 31, Created)
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=LMJ_0712;
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Friedlin;
RA Ivens A.C., Peacock C.S., Worthey E.A., Murphy L., Aggarwal G.,
RA Herriman M., Sisk E., Rajandream M.-A., Aert R., Anupama A.,
RA Apostolou Z., Attipoe P., Bason N., Bauser C., Beck A., Beverly S.M.,
RA Bianchetti G., Borzym K., Bothe G., Bruschi C.V., Collins M.,
RA Cadag E., Chiarioni L., Clayton C., Coulson R.M.R., Cronin A.,
RA Cruz A.K., Davies R.M., De Gaudenzi J., Dobson A.C., Fraser A., Fuchs M., Gabel C.,
RA Fazzelina G., Fosker N., Frasc A.C., Hertz-Fowler C., Hilbert H., Horn D.,
RA Goble A., Goffeau A., Harris D., Knights A., Kube M., Larke N., Litvin L.,
RA Huang Y., Klages S., Knights A., Kube M., Larke N., Litvin L.,
RA Lord A., Louie T., Marra M., Masuy D., Matthews K., Michaeli S.,
RA Mottram J.C., Muller-Auer S., Munden H., Nelson S., Norbertczak H.,
RA Oliver K., O'Neill S., Pentony M., Pohl T.M., Price C., Purnelle B.,
RA Quail M.A., Rabinowitsch E., Reinhardt R., Rieger M., Rinta J.,
RA Robben J., Robertson L., Ruiz J.C., Rutter S., Saunders D.,
RA Schafer M., Schein J., Schwartz D.C., Seeger K., Seyler A., Sharp S.,
RA Shin H., Sivam D., Squares R., Squares S., Tosato V., Vogt C.,
RA Volckaert G., Wambutt R., Warren T., Wedler H., Woodward J., Zhou S.,
RA Zimmermann W., Smith D.F., Blackwell J.M., Stuart K.D., Barrell B.,
RA Myler P.J.;
RT "The genome of the kinetoplastid parasite, Leishmania major.";
RL Science 0:0-0(2005).
DR EMBL; CP000080; AA209527.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 195 AA; E31A79FCBB504D5 CRC64;

Query Match 100.0%; Score 40; DB 2; Length 195;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
Db 157 GGIGDGG 163

RESULT 2
Q9RSI7 DEIRA
ID Q9RSI7 DEIRA PRELIMINARY; PRT; 256 AA.
AC Q9RSI7;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Hypothetical protein DR2137.
GN OrderedLocusNames=DR2137;
OS Deinococcus radiodurans.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=R1 / ATCC 13939 / DSM 20539 / NCIB 9279;
RX MEDLINE=20036896; PubMed=10567266; DOI=10.1126/science.286.5444.1571;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Hatt D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L.A., Utterback T.R., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S.L., Smith H.O., Venter J.C.,
```

```
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1.";
RL Science 286:1571-1577(1999).
DR EMBL; AB02048; AAF11691.1; -; Genomic_DNA.
DR PIR; F75310; F75310.
DR TIGR; DR2137; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 256 AA; 28200 MW; 295746BDBFF36355 CRC64;

Query Match 100.0%; Score 40; DB 2; Length 256;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
Db 154 GGIGDGG 160

RESULT 3
Q7BTU0 YERPE
ID Q7BTU0 YERPE PRELIMINARY; PRT; 289 AA.
AC Q7BTU0;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE DNA polymerase I.
OS Versinia pestis I.
OG Plasmid pMT1.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Versinia.
OX NCBI_TaxID=187410;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=KIM;
RX MEDLINE=98422474; PubMed=9748454;
RA Hu P., Elliott J., McCready P., Skowronski E., Garnes J.,
RA Kobayashi A., Brubaker R.R., Garcia E.;
RT "Structural organization of virulence-associated plasmids of Versinia
RT pestis.";
RL J. Bacteriol. 180:5192-5202(1998).
DR EMBL; AF053947; AAC13209.1; -; Genomic DNA.
DR GO; GO:0008409; F:5'-3' exonuclease activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR002421; 5_3_exonuclease.
DR InterPro; IPR000513; Exo_N_I.
DR InterPro; IPR008918; Hnh2_.
DR Pfam; PF01367; 5_3_exonuc_1.
DR Pfam; PF02739; 5_3_exonuc_N; 1.
DR SMART; SM00475; 53EXOC; 1.
DR SMART; SM00279; Hnh2; 1.
KW Plasmid.
SQ SEQUENCE 289 AA; 32751 MW; 2ADF905515D2B3A5 CRC64;

Query Match 100.0%; Score 40; DB 2; Length 289;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
Db 171 GGIGDGG 177

RESULT 4
Q935F2 SALT1
ID Q935F2 SALT1 PRELIMINARY; PRT; 314 AA.
AC Q935F2;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Putative exonuclease.
GN OrderedLocusNames=HCM2.0010c;
```

OS *Salmonella typhi*.
 OG Plasmid pHCW2.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxID=601;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=CT18;
 RX MEDLINE=21534947; PubMed=11677608; DOI=10.1038/35101607;
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
 RA Churche C.M., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
 RA Felwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
 RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
 RA Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,
 RA Whitehead S., Barrrell B.G.;
 RT "Complete genome sequence of a multiple drug resistant Salmonella
 RT enterica serovar Typhi CT18.";
 RL Nature 413:848-852(2001).
 DR EMBL; AL513384; CAD09877.1; -; Genomic DNA.
 DR GO; GO:0008409; F:5'-3' exonuclease activity; IEA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0016787; F:hydrolase activity; IEA.
 DR InterPro; IPR002421; 5_3_exonuclease.
 DR InterPro; IPR000513; Exo_N_1.
 DR InterPro; IPR008918; HhH2.
 DR Pfam; PF01367; 5_3_exonuc; 1.
 DR Pfam; PF02739; 5_3_exonuc; 1.
 DR SMART; SM00475; 53EXOC; 1.
 DR SMART; SM00279; HhH2; 1.
 KW Complete proteome; Exonuclease; Plasmid.
 SQ SEQUENCE 314 AA; 35755 MW; 4E4B50C355FEA2C8 CRC64;
 Query Match 100.0%; Score 40; DB 2; Length 314;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGIGDGG 7
 Db 196 GGIGDGG 202
 RESULT 5
 ID Q9RIC9 YERPE PRELIMINARY; PRT; 314 AA.
 AC Q9RIC9_074YK2;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
 DE Putative exonuclease (Hypothetical protein).
 GN OrderedLocusNames=YPMW1.51c, pMT068;
 OS *Yersinia pestis*.
 OG Plasmid pMT1.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Yersinia.
 OX NCBI_TaxID=632;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=CO-92 / Biovar Orientalis; PLASMID=pMT1;
 RX MEDLINE=21470413; PubMed=11586360; DOI=10.1038/35097083;
 RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
 RA Prentice M.B., Sebahia M., James K.D., Churcher C.M., Mungall K.L.,
 RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.-M.,
 RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
 RA Felwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
 RA Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.M.,
 RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrrell B.G.;
 RT "Genome sequence of *Yersinia pestis*, the causative agent of plague.";
 RL Nature 413:523-527(2001).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=91001; PLASMID=pMT1;
 Query Match 100.0%; Score 40; DB 2; Length 314;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGIGDGG 7
 Db 196 GGIGDGG 202

RX PubMed=1536893;
 RA Song Y., Tong Z., Wang J., Wang L., Guo Z., Han Y., Zhang J., Pei D.,
 RA Zhou D., Qin H., Pang X., Han Y., Zhai J., Li M., Cui B., Qi Z.,
 RA Jin L., Dai R., Chen F., Li S., Ye C., Du Z., Lin W., Wang J., Yu J.,
 RA Yang H., Wang J., Huang P., Yang R.;
 RT "Complete genome sequence of *Yersinia pestis* strain 91001, an isolate
 RT avirulent to humans.";
 RL DNA Res. 11:179-197(2004).
 DR EMBL; AL117211; CAB5233.1; -; Genomic DNA.
 DR EMBL; AE017045; AAS58701.1; -; Genomic DNA.
 DR PIR; T14694; T14694.
 DR GO; GO:0008409; F:5'-3' exonuclease activity; IEA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0016787; F:hydrolase activity; IEA.
 DR InterPro; IPR002421; 5_3_exonuclease.
 DR InterPro; IPR000513; Exo_N_1.
 DR InterPro; IPR008918; HhH2.
 DR Pfam; PF01367; 5_3_exonuc; 1.
 DR Pfam; PF02739; 5_3_exonuc; 1.
 DR SMART; SM00475; 53EXOC; 1.
 DR SMART; SM00279; HhH2; 1.
 KW Complete proteome; Exonuclease; Hypothetical protein; Plasmid.
 SQ SEQUENCE 314 AA; 35664 MW; 0070BFF6DFB552C7 CRC64;
 Query Match 100.0%; Score 40; DB 2; Length 314;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGIGDGG 7
 Db 196 GGIGDGG 202
 RESULT 6
 ID O68765 YERPE PRELIMINARY; PRT; 331 AA.
 AC O68765;
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)
 DT 01-AUG-1999 (TrEMBLrel. 11, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hypothetical protein Y111.
 GN Name=Y111;
 OS *Yersinia pestis* KIM.
 OG Plasmid pMT-1.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Yersinia.
 OX NCBI_TaxID=187410;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=KIM10+;
 RX MEDLINE=99043898; PubMed=9826348;
 RA Lindler L.E., Plano G.V., Burland V., Mayhew G.F., Blattner F.R.;
 RT "Complete DNA sequence and detailed analysis of the *Yersinia pestis*
 RT KIM5 plasmid encoding murine toxin and capsular antigen.";
 RL Infect. Immun. 66:5731-5742(1998).
 DR EMBL; AF074611; AAC82769.1; -; Genomic DNA.
 DR PIR; T14694; T14694.
 DR PIR; T15026; T15026.
 DR GO; GO:0008409; F:5'-3' exonuclease activity; IEA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0016787; F:hydrolase activity; IEA.
 DR InterPro; IPR002421; 5_3_exonuclease.
 DR InterPro; IPR000513; Exo_N_1.
 DR InterPro; IPR008918; HhH2.
 DR Pfam; PF01367; 5_3_exonuc; 1.
 DR Pfam; PF02739; 5_3_exonuc; 1.
 DR SMART; SM00475; 53EXOC; 1.
 DR SMART; SM00279; HhH2; 1.
 KW Hypothetical protein; Plasmid.
 SQ SEQUENCE 331 AA; 37395 MW; 394B3DEE749D260C CRC64;
 Query Match 100.0%; Score 40; DB 2; Length 331;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
Db 213 GGIGDGG 219

RESULT 7
Q65AK6 YERPE
ID Q65AK6_YERPE PRELIMINARY; PRT; 331 AA.
AC Q65AK6_YERPE
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
OS Yersinia pestis.
DE Yersinia pestis.
OG Plasmid pG8786.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]

NUCLEOTIDE SEQUENCE.
RP PubMed=15385458; DOI=10.1128/IAI.72.10.5613-5621.2004;
RX Golubov A., Neubauer H., Nolting C., Heesemann J., Rakin A.,
RA "Structural Organization of the pFra Virulence-Associated Plasmid of
RT Rhamnose-Positive Yersinia pestis.";
RL Infect. Immun. 72:5613-5621(2004).
DR EMBL; AJ698720; CAG27467.1; -; Genomic DNA.
DR GO; GO:0008409; F.5'-3' exonuclease activity; IEA.
DR GO; GO:0003677; F.DNA binding; IEA.
DR GO; GO:0016787; F.hydrolyase activity; IEA.
DR InterPro; IPR002421; 5_3_exonuclease.
DR InterPro; IPR000513; Exo_N_1.
DR InterPro; IPR008918; HhZ_1.
DR Pfam; PF01367; 5_3_exonuc; 1.
DR Pfam; PF02739; 5_3_exonuc; 1.
DR SMART; SM00475; 53EXOC; 1.
DR SMART; SM00279; HhZ2_1.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 331 AA; 37395 MW; 394B3DER749D260C CRC64;

Query Match 100.0%; Score 40; DB 2; Length 331;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
Db 213 GGIGDGG 219

RESULT 8
Q43494 LUPAN
ID Q43494_LUPAN PRELIMINARY; PRT; 351 AA.
AC Q43494;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nodulin-45 protein.
GN Nodulin-45;
OS Lupinus angustifolius (Narrow-leaved blue lupin).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Genistaeae; Lupinus.
OX NCBI_TaxID=3871;
RN [1]

NUCLEOTIDE SEQUENCE.
RP TISSUE=Legume root nodule;
RC Farnden K.J.;
RA Submitted (JUN-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z12625; CAA78271.1; -; mRNA.
DR PIR; S70765; S70765.
SQ SEQUENCE 351 AA; 38229 MW; 41ED5183AAD4051B CRC64;

Query Match 100.0%; Score 40; DB 2; Length 351;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
Db 272 GGIGDGG 278

RESULT 9
Q43878 LUPAN
ID Q43878_LUPAN PRELIMINARY; PRT; 351 AA.
AC Q43878;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nodulin-45.
OS Lupinus angustifolius (Narrow-leaved blue lupin).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Genistaeae; Lupinus.
OX NCBI_TaxID=3871;
RN [1]

NUCLEOTIDE SEQUENCE.
RP Rice S.J., Grant M.R., Reynolds P.H.S., Farnden K.J.;
RA "DNA sequence of Nodulin-45 from Lupinus angustifolius.";
RT Plant Sci. 0:0-0(1993).
RN [2]

NUCLEOTIDE SEQUENCE.
RX MEDLINE=95201240; PubMed=7894011;
RA Macknight R.C., Reynolds P.H., Farnden K.J.;
RT "Analysis of the lupin Nodulin-45 promoter: conserved regulatory
RT sequences are important for promoter activity.";
RT Plant Mol. Biol. 27:457-466(1995).
RN [3]

NUCLEOTIDE SEQUENCE.
RP Legeay O.M., Rice S.J., Farnden K.J., Reynolds P.H.S.;
RA Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; L12388; AAA73961.1; -; Genomic_DNA.
DR PIR; S70765; S70765.
SQ SEQUENCE 351 AA; 38229 MW; 456F6D887CDA326 CRC64;

Query Match 100.0%; Score 40; DB 2; Length 351;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
Db 272 GGIGDGG 278

RESULT 10
Q6AL55 DESPS
ID Q6AL55_DESPS PRELIMINARY; PRT; 395 AA.
AC Q6AL55;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Probable pleiotropic regulatory protein.
GN OrderedLocustNames=DF2191;
OS Desulfotalea psychrophila.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfobacteriales;
OC Desulfobulbaceae; Desulfotalea.
OX NCBI_TaxID=84980;
RN [1]

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=LSV54 / DSM 12343;
RX PubMed=15305914; DOI=10.1111/j.1462-2920.2004.00665.x;
RA Rabus R., Ruepp A., Frickey T., Rattei T., Fartmann B., Stark M.,
RA Bauer M., Zibat A., Lombardot T., Becker I., Amann J., Gellner K.,
RA Teeling H., Leuschner W.D., Gloeckner F.-O., Lupas A.N., Amann R.,
RA Klenk H.-P.;
RT "The genome of Desulfotalea psychrophila, a sulfate-reducing bacterium
RT from permanently cold Arctic sediments.";
RL Environ. Microbiol. 6:887-902(2004).
DR EMBL; CR522870; CAG36920.1; -; Genomic_DNA.
DR InterPro; IPR000653; Degt DnrJ EryC1.
DR Pfam; PF01041; Degt_DnrJ_EryC1; 1.
KW Complete proteome.
SQ SEQUENCE 395 AA; 43433 MW; E5B4E386988AB47C CRC64;
Query Match 100.0%; Score 40; DB 2; Length 395;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGIGDGG 7
Db 202 GGIGDGG 208
|||||
RESULT 11
O24779 PSEAE
ID O24779 PSEAE PRELIMINARY; PRT; 425 AA.
AC O24779;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE OprE3.
GN Name=oprQ;
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PA01;
RX MEDLINE=99268523; PubMed=10338201;
RA Okamoto K., Gotoh N., Teujimoto H., Yamada H., Yoshihara E., Nakae T.,
RA Nishino T.;
RT "Molecular cloning and characterization of the oprQ gene coding for
RT outer membrane protein OprE3 of Pseudomonas aeruginosa.";
RL Microbiol. Immunol. 43:297-301(1999).
DR EMBL; AB066797; BAA22267.1; -; Genomic_DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0015288; F:porin activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005318; OprD.
DR Pfam; PF03573; OprD; 1.
FT CHAIN 22 425
SQ SEQUENCE 425 AA; 46836 MW; 324012F7316B2EE4 CRC64;
Query Match 100.0%; Score 40; DB 2; Length 425;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGIGDGG 7
Db 305 GGIGDGG 311
|||||
RESULT 12
Q91083 PSEAE
ID Q91083 PSEAE PRELIMINARY; PRT; 425 AA.
AC Q91083;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Probable outer membrane protein.

GN OrderedLocNames=PA2760;
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043; DOI=10.1038/35023079;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H. Jr., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
DR EMBL; AE004704; AAG06148.1; -; Genomic_DNA.
DR PIR; G83299; G83299.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0015288; F:porin activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005318; OprD.
DR Pfam; PF03573; OprD; 1.
KW Complete proteome.
SQ SEQUENCE 425 AA; 46850 MW; 9224A9B5EA4B2EE0 CRC64;
Query Match 100.0%; Score 40; DB 2; Length 425;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGIGDGG 7
Db 305 GGIGDGG 311
|||||
RESULT 13
Q5QQ59 PSEFL
ID Q5QQ59 PSEFL PRELIMINARY; PRT; 426 AA.
AC Q5QQ59;
DT 01-FEB-2005 (TREMBLrel. 29, Created)
DT 01-FEB-2005 (TREMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TREMBLrel. 29, Last annotation update)
DE Outer membrane protein, OprE3 precursor (Fragment).
GN Name=oprQ;
OS Pseudomonas fluorescens.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=294;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MF37;
RA Jaouen T., Chevalier S., Coquet L., Marvin-Guy L., Orange N., De E.;
RT "Functional characterization of two outer membrane proteins from
RT Pseudomonas fluorescens: OprE1 and OprE3.";
RN Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
[2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MF37;
RA Jaouen T.;
RT "Caractérisations structurales et fonctionnelles de protéines de
RT membrane externe de Pseudomonas psychrotrophes et mesophiles.";
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ866544; CAI28815.1; -; Genomic_DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0015288; F:porin activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005318; OprD.
DR Pfam; PF03573; OprD; 1.
KW Signal.
FT SIGNAL 1 21 Potential.
FT CHAIN 22 >426 outer membrane protein, OprE3.

```
FT NON_TER 426 426
SQ SEQUENCE 426 AA; 46224 MW; 2F2F4E683E963B87 CRC64;

Query Match 100.0%; Score 40; DB 2; Length 426;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
Db 317 GGIGDGG 323

RESULT 14
DRTS_TRYBB
ID DRTS_TRYBB STANDARD; PRT; 527 AA.
AC Q27783;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Bifunctional dihydrofolate reductase-thymidylate synthase (DHFR-TS)
DE [Includes: Dihydrofolate reductase (EC 1.5.1.3); Thymidylate synthase
DE (EC 2.1.1.45)].
OS Trypanosoma brucei.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=427;
RX MEDLINE=96089823; PubMed=8538681; DOI=10.1016/0166-6851(95)00059-A;
RA Gamarro F., Yu P.L., Zhao J., Eaman U., Greene P.J., Santi D.;
RT "Trypanosoma brucei dihydrofolate reductase-thymidylate synthase: gene
RT isolation and expression and characterization of the enzyme.";
RL Mol. Biochem. Parasitol. 72:11-22(1995).
CC -!- CATALYTIC ACTIVITY: 5,10-methylenetetrahydrofolate + dUMP =
CC dihydrofolate + dUMP.
CC -!- CATALYTIC ACTIVITY: 5,6,7,8-tetrahydrofolate + NADP(+) = 7,8-
CC dihydrofolate + NADPH.
CC -!- PATHWAY: Deoxyribonucleotide biosynthesis.
CC -!- PATHWAY: Cofactor biosynthesis; tetrahydrofolate biosynthesis;
CC tetrahydrofolate from 2-amino-4-hydroxy-6-hydroxymethyl-7,8-
CC dihydropteridine diphosphate and 4-aminobenzoate: step 3 [final
CC step].
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SIMILARITY: In the N-terminal section; belongs to the
CC dihydrofolate reductase family.
CC -!- SIMILARITY: In the C-terminal section; belongs to the thymidylate
CC synthase family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; U20781; AAA91362.1; -; Genomic_DNA.
CC HSSP; P13100; 1CI7.
CC InterPro; IPR012262; DHFR-TS.
CC InterPro; IPR001796; DHFR reg.
CC InterPro; IPR000398; Thymidylat_synth.
CC Pfam; PF00186; DHFR 1; 1.
CC Pfam; PF00303; Thymidylat synt; 1.
CC PIRSF; PIRSF000389; DHFR-TS; 1.
CC PRINTS; PR00070; DHFR.
CC PRINTS; PR00108; THYMSINTHASE.
CC ProDom; PD00180; Thymidylat_synth; 1.
CC PROSITE; PS00075; DHFR; 1.
CC PROSITE; PS00091; THYMIDYLATE SYNTHASE; 1.
CC Methyltransferase; Multifunctional enzyme; NADP;
CC Nucleotide biosynthesis; One-carbon metabolism; Oxidoreductase;
CC Transferase.
CC REGION 1 242 Dihydrofolate reductase.
CC REGION 243 527 Thymidylate synthase.
CC ACT_SITE 409 409 By similarity.
```

```
SQ SEQUENCE 527 AA; 58805 MW; 39982FC53BF7601 CRC64;

Query Match 100.0%; Score 40; DB 1; Length 527;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
Db 39 GGIGDGG 45

RESULT 15
Q582G3_9TRYP PRELIMINARY; PRT; 527 AA.
AC Q582G3;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Dihydrofolate reductase-thymidylate synthase (EC 1.5.1.3).
DE ORFNames=TB927.7.5480;
GN Trypanosoma brucei.
OS Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5691;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GUTat10.1;
RA Ghedin E., Blandin G., Bartholomeu D., Caler E., Haas B., Hannick L.,
RA Shallom J., Hou L., Djikeng A., Feldblyum T., Hostetler J.,
RA Johnson J., Jones K., Koo H.L., Larkin C., Pai G., Peterson J.,
RA Khalak H.G., Salzberg S., Simpson A.J., Tallon L., Van Aken S.,
RA Wanless D., White O., Wortman J., Fraser C.M., El-Sayed N.M.A.;
RA Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RL [2]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=GUTat10.1;
RC El-Sayed N.M., Khalak H., Adams M.D.;
RA Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RL [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GUTat10.1;
RA Haas B., Blandin G., El-Sayed N.;
RA Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
RL EMBL; AC091553; AAX78868.1; -; Genomic_DNA.
KW Oxidoreductase.
SQ SEQUENCE 527 AA; 58850 MW; 1917FE451F6A9EDA CRC64;

Query Match 100.0%; Score 40; DB 2; Length 527;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
Db 39 GGIGDGG 45

RESULT 16
DNM1L_MOUSE
ID DNM1L_MOUSE STANDARD; PRT; 742 AA.
AC Q8K1M6; Q8BNQ5; Q8BG64; Q8CGD0; Q8K1A1;
DT 10-MAY-2005 (Rel. 47, Created)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Dynamin 1-like protein (EC 3.6.5.5) (dynamin-related protein 1)
DE (Dynamin family member proline-rich carboxyl-terminal domain less
DE (Dyple)).
DE Name=Dnm1l; Synonyms=Drp1;
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muroidae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 2).
```

RC TISSUE=Osteoclast; DOI=10.1016/j.bbrc.2003.10.008;
RX PubMed=14592431;
RA Honda S.; Hirose S.;
RT "Stage-specific enhanced expression of mitochondrial fusion and
RL fission factors during spermatogenesis in rat testis.";
RN Biochem. Biophys. Res. Commun. 311:424-432(2003).
[2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 3).
RC STRAIN=C57BL/6J; TISSUE=Adipose tissue, and Spinal ganglion;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaide I., Oeato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojohori T.,
RA Baldarelli R., Hill D.P., Bult C.P., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Girardot S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasaki Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramchandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita K.,
RA Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Haya A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL 60,770 full-length cDNAs";
RN Nature 420:563-573(2002).
[3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORMS 3 AND 4).
RC STRAIN=C57BL/6; TISSUE=Brain, Mammary gland, and thymus;
RX MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Rosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs S.J.,
RA Fahy J., Helton E., Kettner M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.B.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 448-659 (ISOFORM 1).
RX PubMed=11544199; DOI=10.1101/gr.185501;
RA Piao Y., Ko N.T., Lim M.K., Ko M.S.H.;
RT "Construction of long-transcript enriched cDNA libraries from
RT submicrogram amounts of total RNAs by a universal PCR amplification
RL method";
RN Genome Res. 11:1553-1558(2001).
[5]
RP TISSUE SPECIFICITY.

EX PubMed=9422767; DOI=10.1074/jbc.273.2.1044;
RA Kamimoto T., Nagai Y., Onogi H., Muro Y., Wakabayashi T., Hagiwara M.;
RT "Dymple, a novel dynamin-like high molecular weight GTPase lacking a
RL proline-rich carboxyl-terminal domain in mammalian cells.";
RN J. Biol. Chem. 273:1044-1051(1998).
CC -!- FUNCTION: Functions in mitochondrial and peroxisomal division
CC probably by regulating membrane fission. Enzyme hydrolyzing GTP
CC that oligomerizes to form ring-like structures and is able to
CC remodel membranes. May also play a role on organelles of the
CC secretory pathway (By similarity).
CC -!- CATALYTIC ACTIVITY: GTP + H₂O = GDP + phosphate.
CC -!- SUBUNIT: Homotrimer; N-terminal part binds to the C-terminal
CC part of another DNML. Can self-assemble in multimeric ring-like
CC structures. Interacts with GSK3B and TTC11 (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic; mainly cytosolic. Also
CC membrane-associated. Localizes to mitochondria at spots of
CC division events. Associated with peroxisomal membranes it is
CC recruited in part by PEX11B. May also be associated with
CC endoplasmic reticulum tubules and cytoplasmic vesicles and found
CC to be perinuclear (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=5;
CC Name=1;
CC IsoId=Q8K1M6-1; Sequence=Displayed;
CC Note=No experimental confirmation available;
CC Name=2;
CC IsoId=Q8K1M6-2; Sequence=VSP_013695;
CC Name=3;
CC IsoId=Q8K1M6-3; Sequence=VSP_013690, VSP_013694;
CC Name=4;
CC IsoId=Q8K1M6-4; Sequence=VSP_013689, VSP_013691, VSP_013695;
CC Name=5;
CC IsoId=Q8K1M6-5; Sequence=VSP_013690, VSP_013692, VSP_013693;
CC -!- TISSUE SPECIFICITY: Expressed in the cerebellum and in several
CC regions of the cerebrum and diencephalon. Strongly expressed in
CC the cerebellar Purkinje cells and in the pontile giant neurons.
CC -!- PTM: Phosphorylated by GSK3B (By similarity).
CC -!- SIMILARITY: Belongs to the dynamin family.
CC -!- SIMILARITY: Contains 1 GED domain.
CC -!- CAUTION: This is a conceptual translation (isoform 1).

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

DR EMBL; AB079133; BAC06576.1; -; mRNA.
DR EMBL; AK051443; BAC34640.1; -; mRNA.
DR EMBL; AK080871; BAC38054.1; -; mRNA.
DR EMBL; BC027538; AAH27538.1; -; mRNA.
DR EMBL; BC040777; AAH40777.1; -; mRNA.
DR EMBL; BC079635; AAH79635.1; -; mRNA.
DR EMBL; CF914619; -; NOT ANNOTATED CDS; mRNA.
DR Ensembl; ENSMUSG0000022789; Mus musculus.
DR MGI; MGI:1921256; Dm11.
DR InterPro; IPR001401; Dynamin.
DR InterPro; IPR000375; Dynamin_central.
DR InterPro; IPR003130; GED.
DR Pfam; PF01031; Dynamin_M; 1.
DR Pfam; PF00350; Dynamin_N; 1.
DR Pfam; PF02212; GED; 1.
DR PRINTS; PR00195; DYNAMIN.
DR SMART; SM00053; DYNC; 1.
DR SMART; SM00302; GED; 1.
DR PROSITE; PS00410; DYNAMIN; 1.
KW Alternative splicing; GTP-binding; Hydrolase; Membrane;
KW Nucleotide-binding; phosphorylation.
FT DOMAIN 645 736 GED.
FT NP_BIND 32 39 GTP (By similarity).
FT NP_BIND 152 156 GTP (By similarity).
FT NP_BIND 221 224 GTP (By similarity).
FT REGION 1 349 N-terminal dimerization domain (By

FT REGION 454 654 similarity).

FT REGION 548 742 Interaction with GSK3B (By similarity).

FT VARSPLIC 1 104 C-terminal dimerization domain (By similarity).

FT VARSPLIC 85 90 Missing (in isoform 4).

FT VARSPLIC 105 105 /FTId=VSP_013689.

FT VARSPLIC 214 227 Missing (in isoform 3 and isoform 5).

FT VARSPLIC 228 742 /FTId=VSP_013690.

FT VARSPLIC 539 575 K -> M (in isoform 4).

FT VARSPLIC 539 564 /FTId=VSP_013691.

FT CONFLICT 165 165 RRTLAIVTKLDLMD -> KGRCLYLMVDVLDQ (in isoform 5).

FT CONFLICT 320 320 /FTId=VSP_013692.

FT CONFLICT 519 519 Missing (in isoform 5).

FT SEQUENCE 742 AA; 82658 MW; 5A8679B61A444847 CRC64;

Query Match 100.0%; Score 40; DB 1; Length 742;

Best Local Similarity 100.0%; Pred. No. 3.3e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7

Db 581 GGIGDGG 587

RESULT 17

DNMIL RAT STANDARD; PRT: 755 AA.

AC Q3503; Q35318; Q35319; Q35320; Q35321; Q35322; Q35323; Q5U2W1;

AC Q792T7; Q9R234; Q9R277;

DT 10-MAY-2005 (Rel. 47, Created)

DT 10-MAY-2005 (Rel. 47, Last sequence update)

DT 13-SEP-2005 (Rel. 48, Last annotation update)

DE Dynamin 1-like protein (RC 3.6.5.5) (Dynamin-like protein).

GN Name=Dnm1; Synonyms=Dlp1;

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

OC Muroidae; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORMS 1; 2; 3; 4; 5 AND 6), PROTEIN

RP SEQUENCE OF 39-48; 62-74; 113-126; 147-160; 174-180 AND 700-710,

RP SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.

RC TISSUE=Brain; and Liver;

RX PubMed=9472031; DOI=10.1083/jcb.140.4.779;

RA Yoon Y., Pitts K.R., Dahan S., McNiven M.A.;

RT "A novel dynamin-like protein associates with cytoplasmic vesicles and

RT tubules of the endoplasmic reticulum in mammalian cells.";

RL J. Cell Biol. 140:779-793(1998).

RN [2]

RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORMS 1; 3 AND 5), TISSUE SPECIFICITY,

RP INTERACTION WITH GSK3B, AND PHOSPHORYLATION BY GSK3B.

RC TISSUE=Brain;

RX MEDLINE=20145530; PubMed=10679301; DOI=10.1006/bbrc.2000.2197;

RA Chen C.-H., Hwang S.-L., Howng S.-L., Chou C.-K., Hong Y.-R.;

RT "Three rat brain alternative splicing dynamin-like protein variants:

RT interaction with the glycogen synthase kinase 3beta and action as a

RT substrate.";

RL Biochem. Biophys. Res. Commun. 268:893-898(2000).

RN [3]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 2).

RC TISSUE=Brain;

RX NIH - Mammalian Gene Collection (MGC) project;

RG Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.

RN [4]

RP MUTAGENESIS OF LYS-38 AND ASP-231, AND OLIGOMERIZATION.

RX MEDLINE=21437878; PubMed=11553726;

RA Yoon Y., Pitts K.R., McNiven M.A.;

RT "Mammalian dynamin-like protein Dlp1 tubulates membranes.";

RL Mol. Biol. Cell 12:2894-2905(2001).

RN [5]

RP SUBCELLULAR LOCATION, INDUCTION BY BEZAFIBRATE, AND FUNCTION.

RX MEDLINE=22499627; PubMed=12499366; DOI=10.1074/jbc.M211761200;

RA Koch A., Thiemann M., Grabenbauer M., Yoon Y., McNiven M.A.,

RA Schrader M.;

RT "Dynamin-like protein 1 is involved in peroxisomal fission.";

RL J. Biol. Chem. 278:8597-8605(2003).

RN [6]

RP FUNCTION, AND INTERACTION WITH TTC11.

RX MEDLINE=22745481; PubMed=12861026;

RX DOI=10.1128/MCB.23.15.5409-5420.2003;

RA Yoon Y., Krueger E.W., Oswald B.J., McNiven M.A.;

RT "The mitochondrial protein hFis1 regulates mitochondrial fission in

RT mammalian cells through an interaction with the dynamin-like protein

RT Dlp1.";

RL Mol. Cell. Biol. 23:5409-5420(2003).

CC -!- FUNCTION: Functions in mitochondrial and peroxisomal division

CC probably by regulating membrane fission. Enzyme hydrolyzing GTP

CC that oligomerizes to form ring-like structures and is able to

CC remodel membranes. May also play a role on organelles of the

CC secretory pathway.

CC -!- CATALYTIC ACTIVITY: GTP + H(2)O = GDP + phosphate.

CC -!- SUBUNIT: Homotetramer; N-terminal part binds to the C-terminal

CC part of another DNM1L. Can self-assemble in multimeric ring-like

CC structures. Interacts with GSK3B and TTC11.

CC -!- SUBCELLULAR LOCATION: Cytoplasmic; mainly cytosolic. Also

CC membrane-associated. Localizes to mitochondria at spots of

CC division events. Associated with peroxisomal-membranes it is

CC recruited in part by PEX11B. May also be associated with

CC endoplasmic reticulum tubules and cytoplasmic vesicles and found

CC to be perinuclear (By similarity).

CC -!- ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=6;

CC Name=1;

CC IsoId=O35303-1; Sequence=Displayed;

CC Name=2;

CC IsoId=O35303-2; Sequence=VSP_013697, VSP_013701;

CC Name=3; Synonyms=DLP1-37;

CC IsoId=O35303-3; Sequence=VSP_013698, VSP_013699;

CC Name=4;

CC IsoId=O35303-4; Sequence=VSP_013696;

CC Note=No experimental confirmation available;

CC Name=5; Synonyms=DLP1-11;

CC IsoId=O35303-5; Sequence=VSP_013702;

CC Name=6;

CC IsoId=O35303-6; Sequence=VSP_013700;

CC Note=No experimental confirmation available;

CC -!- TISSUE SPECIFICITY: Expressed in all tissues tested (at protein

CC level). Longer isoforms are preferentially expressed in brain.

CC -!- INDUCTION: By bezafibrate.

CC -!- PTM: Phosphorylated by GSK3B.

CC -!- SIMILARITY: Belongs to the dynamin family.

CC -!- SIMILARITY: Contains 1 GED domain.

CC -----

CC This Swiss-Prot entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use as long as its content is in no way modified and this statement is not

CC removed.

CC -----

DR EMBL; AF019043; AAB72197.1; -; mRNA.

DR EMBL; AF020207; AAB71232.1; -; mRNA.

DR EMBL; AF020208; AAB71233.1; -; mRNA.

DR EMBL; AF020209; AAB71234.1; -; mRNA.

DR EMBL; AF020210; AAB71235.1; -; mRNA.

DR EMBL; AF020211; AAB71236.1; -; mRNA.

DR EMBL; AF020212; AAB71237.1; -; mRNA.

```

DR EMBL; AF020213; AAB71238.1; -; mRNA.
DR EMBL; AF107048; AAD22412.1; -; mRNA.
DR EMBL; AF132727; AAD31278.1; -; mRNA.
DR EMBL; BC085843; AAH85843.1; -; mRNA.
DR ENSEMBL; ENSRNOG00000001813; Rattus norvegicus.
DR RGD; 620416; Dmml.
DR InterPro; IPR001401; Dynamain.
DR InterPro; IPR000375; Dynamain_central.
DR InterPro; IPR003130; GED.
DR Pfam; PF01031; Dynamain_N; 1.
DR Pfam; PF00350; Dynamain_N; 1.
DR Pfam; PF02212; GED; 1.
DR PRINTS; PR00195; DYNAMIN.
DR SMART; SM00053; DYNC; 1.
DR SMART; SM00302; GED; 1.
DR PROSITE; PS00410; DYNAMIN; 1.
KW Alternative splicing; Direct protein sequencing; GTP-binding;
KW Hydrolase; Membrane; Nucleotide-binding; Phosphorylation.
FT DOMAIN 658 749
FT NP_BIND 32 39
FT NP_BIND 159 163
FT NP_BIND 228 231
FT REGION 1 356
FT REGION 461 667
FT REGION 555 755
FT VARSPLIC 84 97
FT VARSPLIC 84 96
FT VARSPLIC 84 84
FT VARSPLIC 546 596
FT VARSPLIC 546 582
FT VARSPLIC 546 571
FT VARSPLIC 572 582
FT MUTAGEN 38 38
FT MUTAGEN 231 231
FT CONFLICT 517 517
FT CONFLICT 600 600
SQ SEQUENCE 755 AA; 83908 MW; 056835390794C43 CRC64;

Query Match 100.0%; Score 40; DB 1; Length 755;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGIGDGG 7
Db 588 GGIGDGG 594

RESULT 18
Q4NTR7_9DELT PRELIMINARY; PRT; 984 AA.
AC Q4NTR7;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Protein kinase.
GN ORFNames=AdedRAFT_2163;
OS Anaeromyxobacter dehalogenans 2CP-C.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;
OC Cytophacterales; Myxococcaceae; Anaeromyxobacter.
OX NCBI_taxid=290397;
RN [1]

```

```

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=2CP-C;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RA Hammon N., Israni S., Pittluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Anaeromyxobacter
RT dehalogenans 2CP-C.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=2CP-C;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Anaeromyxobacter
RT dehalogenans 2CP-C.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DDBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAHD01000016; EAL79048.1; -; Genomic_DNA.
KW Kinase.
SQ SEQUENCE 984 AA; 105504 MW; 048AD4B22E6C994A CRC64;

Query Match 100.0%; Score 40; DB 2; Length 984;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGIGDGG 7
Db 229 GGIGDGG 235

RESULT 19
NIT4_NEUCR
ID NIT4_NEUCR STANDARD; PRT; 1090 AA.
AC P28349; Q7RVG9;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Nitrogen assimilation transcription factor nit-4.
GN Name=nit-4; ORFNames=NCU08294;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_taxid=5141;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RX MEDLINE=92017855; PubMed=1840634;
RA Yuan G.-F., Fu Y.-H., Marzluf G.A.;
RT "nit-4, a pathway-specific regulatory gene of Neurospora crassa,
RT encodes a protein with a putative binuclear zinc DNA-binding domain.";
RL Mol. Cell. Biol. 11:5735-5745(1991).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92149315; PubMed=1531376;
RA Yuan G.-F., Marzluf G.A.;
RT "Molecular characterization of mutations of nit-4, the pathway-
RT specific regulatory gene which controls nitrate assimilation in
RT Neurospora crassa.";
RL Mol. Microbiol. 6:67-73(1992).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=74-OR23-1A / FGSC 987;
RX MEDLINE=22598136; PubMed=12712197; DOI=10.1038/nature01554;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Lefebvre D., Fittburgh W., Ma L.-J., Smirnov S., Purcell S., Rehm B.,
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Tanakiev P., Bell-Pedersen D., Nelson M.A.,
RA Werner-Washburne M., Selitrennikoff C.P., Kinsey J.A., Braun E.L.,
RA Zelter A., Schulte U., Kothe G.O., Jedd G., Mewes H.-W., Staben C.,
RA Marcotte E., Greenberg D., Roy A., Foley K., Naylor J.,
RA Stange-Thomann N., Barrett R., Gnerre S., Kamal M., Kamvysselis M.,
RA Mauceli E., Bielke C., Rudd S., Frishman D., Krystofova S.,

```

```
RA Rasmussen C., Metzberg R.L., Perkins D.D., Kroken S., Cogoni C.,
RA Macino G., Catcheside D.E.A., Li W., Pratt R.J., Omani S.A.,
RA DeSouza C.P.C., Glass N.L., Orbach M.J., Berglund J.A., Voelker R.,
RA Yarden O., Plamann M., Seiler S., Dunlap J.C., Radford A., Aramayo R.,
RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbola D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.W.;
RT "The genome sequence of the filamentous fungus Neurospora crassa.";
RL Nature 422:859-868(2003).
CC -!- FUNCTION: Pathway-specific regulatory gene of nitrate
CC assimilation; it activates the transcription of the genes for
CC nitrate and nitrite reductases.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- DOMAIN: The glutamine-rich domain might function in activating
CC gene expression.
CC -!- SIMILARITY: Contains 1 Zn(2)-C6 fungal-type DNA-binding domain.
CC
CC This Swiss-Prot entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC -----
CC EMBL; M80368; AAA33602.1; -; Genomic DNA.
CC EMBL; AABX01000206; EAA33271.1; -; Genomic DNA.
CC PIR; A41696; A41696.
CC HSSP; P07272; 1PY1.
CC
CC TRANSFAC; T02845; -.
CC InterPro; IPR007219; Fungal trans.
CC InterPro; IPR001138; Fungi_Trcsrp_N.
CC Pfam; PF04082; Fungal trans; 1.
CC Pfam; PF00172; Zn_Cln5; 1.
CC
CC SMART; SM00066; GAL4; 1.
CC PROSITE; PS00463; ZN2_Cy6_FUNGAL_1; 1.
CC PROSITE; PS50048; ZN2_Cy6_FUNGAL_2; 1.
CC Activator; DNA-binding; Metal-binding; Nitrate assimilation;
KW Nuclear protein; Transcription; Transcription regulation; Zinc.
FT DNA_BIND 53 81 Zn(2)-C6 fungal-type.
FT COMBIAS 121 139 Asp/Glu-rich (acidic).
FT COMBIAS 213 229 Asp/Glu-rich (acidic).
FT COMBIAS 429 450 Asp/Glu-rich (acidic).
FT COMBIAS 752 754 Pro-rich.
FT COMBIAS 755 859 Gln-rich.
FT COMBIAS 992 1024 Poly-Gln.
FT CONFLICT 98 98 K -> KP (in Ref. 1).
FT CONFLICT 276 277 NA -> KP (in Ref. 3).
FT CONFLICT 467 467 L -> S (in Ref. 1).
FT SEQUENCE 1090 AA; 120244 MW; 881D89172EDD6114 CRC64;
SQ
Query Match 100.0%; Score 40; DB 1; Length 1090;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGIGDGG 7
Db 952 GGIGDGG 958
Db
RESULT 20
Q5A4Z5 CANAL PRELIMINARY; PRT; 162 AA.
AC Q5A4Z5;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein.
GN ORFNames=Cao19.8566, Cao19.951;
OS Candida albicans SC5314.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=237561;
[1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SC5314;
```

```
RX PubMed=15123810; DOI=10.1073/pnas.0401648101;
RA Jones T., Federspiel N.A., Chibana H., Dungan J., Kalman S.,
RA Magee B.W., Newport G., Thorstenson Y.R., Agabian N., Magee P.T.,
RA Davis R.W., Scherer S.;
RT "The diploid genome sequence of Candida albicans.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:7329-7334(2004).
[2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SC5314;
RA Dungan J., Kuo A., Newport G., Lan C.-Y., Iijima C., Adegbola O.,
RA Roberts J., Persson K., Donnelly S., Favoreto S., Tzung K.-W.,
RA Jones T., Scherer S., Agabian N.;
RT "Annotation of the Genome of Candida albicans.";
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AACQ01000083; EAK97803.1; -; Genomic DNA.
DR EMBL; AACQ01000082; EAK97864.1; -; Genomic DNA.
KW Hypothetical protein.
SQ SEQUENCE 162 AA; 18284 MW; 54554BB08920F2E7 CRC64;
Query Match 97.5%; Score 39; DB 2; Length 162;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGIGDGG 7
Db 140 GGIGDGG 146
Db
RESULT 21
Q4TID1 TETNG PRELIMINARY; PRT; 261 AA.
AC Q4TID1;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome undetermined SCAP2245, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG00038349001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
[1]
RP NUCLEOTIDE SEQUENCE.
RA Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouteau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dessat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castellani V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crolius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
[2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; CAEE01002245; CAF87351.1; -; Genomic DNA.
DR NON_TER 1
FT SEQUENCE 261 AA; 27539 MW; C9CB72AAEC24A3C2 CRC64;
SQ
Query Match 97.5%; Score 39; DB 2; Length 261;
RC STRAIN=SC5314;
```

Best Local Similarity 85.7%; Pred. No. 1.7e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 6; Conservative 1;

QY 1 GGIGDGG 7
Db 86 GGVGDDG 92

RESULT 22

Q60AS9 METCA
ID Q60AS9 METCA PRELIMINARY; PRT; 294 AA.
AC Q60AS9
DT 25-OCT-2004 (TRENBLrel. 28, Created)
DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TRENBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=MCA0765;
OS Methylococcus capsulatus;
OC Bacteria; Proteobacteria; Gammaproteobacteria; Methylococcales;
OC Methylococcaceae; Methylococcus.
OX NCBI_TaxID=414;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Bath / NCIMB 11132;
RX PubMed=15393840; DOI=10.1371/journal.pbio.0020303;
RA Ward N.B., Larsen O., Sakwa J., Bruset L., Khouri H.M., Durkin A.S.,
RA Dimitrov G., Jiang L., Scanlan D., Kang K.H., Lewis M.R., Nelson K.E.,
RA Methe B.A., Wu M., Heidelberg J.F., Paulsen I.T., Fouts D.E.,
RA Ravel J., Tettelin H., Ren Q., Read T.D., DeBooy R.T., Seshadri R.,
RA Salzberg S.L., Jensen H.B., Birkeland N.K., Nelson W.C., Dodson R.J.,
RA Grindback S.H., Holt I.E., Eidhammer I., Jonassen I., Vanaken S.,
RA Utterback T.R., Feldblyum T.V., Fraser C.M., Lillehaug J.R.,
RA Eisen J.A.;
RT "Genomic insights into methanotrophy: the complete genome sequence of
RT Methylococcus capsulatus (Bath).";
RL PLOS Biol. 2:1616-1628(2004).
DR EMBL; AE017282; AAU92964.1; -; Genomic_DNA.
DR TIGR; MCA0765; -;
DR InterPro; IPR003368; Chlamydia_PMP.
DR Pfam; PF02415; Chlam_PMP; 5.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 294 AA; 29181 MW; 4FAA76CE6F0478C8 CRC64;

Query Match 97.5%; Score 39; DB 2; Length 294;
Best Local Similarity 85.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
Db 133 GGVGDDG 139

RESULT 23

Q7XHX5 ORYSA
ID Q7XHX5 ORYSA PRELIMINARY; PRT; 333 AA.
AC Q7XHX5;
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-FEB-2005 (TRENBLrel. 29, Last annotation update)
DE Hypothetical protein OSUNBA0008J01.9 (Hypothetical protein
DE OJ1205_F02.22).
GN Name=OSJNBa0008J01.9; Synonyms=OJ1205_F02.22;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriactoidae; Oryzaeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC
RT clone:OSJNBa0008J01.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.

QY 1 GGIGDGG 7
Db 86 GGVGDDG 92

Q60AS9 METCA
ID Q60AS9 METCA PRELIMINARY; PRT; 294 AA.
AC Q60AS9
DT 25-OCT-2004 (TRENBLrel. 28, Created)
DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TRENBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=MCA0765;
OS Methylococcus capsulatus;
OC Bacteria; Proteobacteria; Gammaproteobacteria; Methylococcales;
OC Methylococcaceae; Methylococcus.
OX NCBI_TaxID=414;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Bath / NCIMB 11132;
RX PubMed=15393840; DOI=10.1371/journal.pbio.0020303;
RA Ward N.B., Larsen O., Sakwa J., Bruset L., Khouri H.M., Durkin A.S.,
RA Dimitrov G., Jiang L., Scanlan D., Kang K.H., Lewis M.R., Nelson K.E.,
RA Methe B.A., Wu M., Heidelberg J.F., Paulsen I.T., Fouts D.E.,
RA Ravel J., Tettelin H., Ren Q., Read T.D., DeBooy R.T., Seshadri R.,
RA Salzberg S.L., Jensen H.B., Birkeland N.K., Nelson W.C., Dodson R.J.,
RA Grindback S.H., Holt I.E., Eidhammer I., Jonassen I., Vanaken S.,
RA Utterback T.R., Feldblyum T.V., Fraser C.M., Lillehaug J.R.,
RA Eisen J.A.;
RT "Genomic insights into methanotrophy: the complete genome sequence of
RT Methylococcus capsulatus (Bath).";
RL PLOS Biol. 2:1616-1628(2004).
DR EMBL; AE017282; AAU92964.1; -; Genomic_DNA.
DR TIGR; MCA0765; -;
DR InterPro; IPR003368; Chlamydia_PMP.
DR Pfam; PF02415; Chlam_PMP; 5.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 294 AA; 29181 MW; 4FAA76CE6F0478C8 CRC64;

Query Match 97.5%; Score 39; DB 2; Length 333;
Best Local Similarity 85.7%; Pred. No. 2.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
Db 6 GGVGDDG 12

RESULT 24

Q61EN4 ORYSA
ID Q61EN4 ORYSA PRELIMINARY; PRT; 334 AA.
AC Q61EN4
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE WRKY transcription factor 47.
GN Name=WRKY47;
OS Oryza sativa (indica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriactoidae; Oryzaeae; Oryza.
OX NCBI_TaxID=39946;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15047897; DOI=10.1104/pp.103.034967;
RA Zhang Z.-L., Xie Z., Zou X., Casaretto J., Ho T.H., Shen Q.J.;
RT "A rice WRKY gene encodes a transcriptional repressor of the
RT gibberellin signaling pathway in aleurone cells.";
RL Plant Physiol. 134:1500-1513(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Zhang Z.-L., Xie Z., Zou X., Casaretto J., Ho T.-h.D., Shen Q.J.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
CC -!- MISCELLANEOUS: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ third party annotation (TPA) entry.
DR EMBL; BK005050; DAA05112.1; -; Genomic_DNA.
DR Gramene; Q61EN4; -;
DR GO; GO:0003677; F:DNA binding; IEA.
DR InterPro; IPR003657; WRKY.
DR Pfam; PF03106; WRKY; 1.
DR PROSITE; PS0811; WRKY; 1.
SQ SEQUENCE 334 AA; 34963 MW; A9325AACFC93F7C CRC64;

Query Match 97.5%; Score 39; DB 2; Length 334;
Best Local Similarity 85.7%; Pred. No. 2.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
Db 6 GGVGDDG 12

RESULT 25

Q88R66_PSEPK
ID Q88R66_PSEPK PRELIMINARY; PRT; 439 AA.
AC Q88R66;

DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Outer membrane protein OprE3.
 GN Name=oprQ; OrderedLocusNames=PP0268;
 OS Pseudomonas putida (strain KT2440).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=160488;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22423060; PubMed=12534463;
 RX DOI=10.1046/j.1462-2920.2002.00366.x;
 RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
 RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
 RA Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,
 RA Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
 RA Hance I., Chris Lee P., Holtzapple E.K., Scanlan D., Tran K.,
 RA Moazzaz A., Utterback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,
 RA Wedler H., Lauber J., Stjepandic D., Hohnselt J., Straetz M., Heim S.,
 RA Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tuenmiller B.,
 RA Fraser C.M.;
 RT "Complete genome sequence and comparative analysis of the
 RT metabolically versatile Pseudomonas putida KT2440.";
 RL Environ. Microbiol. 4:799-808(2002).
 DR EMBL; AB016775; AA65899.1; -; Genomic_DNA.
 DR TIGR; PP0268; -.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0015288; P:porin activity; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR005318; OprD.
 DR Pfam; PF03573; OprD; 1.
 KW Complete proteome.
 SQ SEQUENCE 439 AA; 4787 MW; 161AB9EC3287B977 CRC64;

 Query Match 97.5%; Score 39; DB 2; Length 439;
 Best Local Similarity 85.7%; Pred. No. 2.9e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

 QY 1 GGIGDGG 7
 DB 311 GGVGDGG 317

 RESULT 26
 Q9NXG7 HUMAN
 ID Q9NXG7 HUMAN PRELIMINARY; PRT; 525 AA.
 AC Q9NXG7;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hypothetical protein FLJ20261.
 GN Name=KRT24;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Colon mucosa;
 RC Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
 RA Suzuki Y., Ohtayashi M., Nishi T., Shibahara T., Tanaka T.,
 RA Nakamura Y., Isegai T., Sugano S.;
 RA Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: Belongs to the intermediate filament family.
 DR EMBL; AK00268; BA91044.1; -; mRNA.
 DR HSP; P08670; ICK7.
 DR Ensembl; ENSG00000167916; Homo sapiens.
 DR HGNC; HGNC:18527; KRT24.
 DR GO; GO:0005882; C:intermediate filament; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR001664; IF.

DR InterPro; IPR002957; Keratin I.
 DR Pfam; PF00038; Filament; 1.
 DR PRINTS; PR01248; TYPE1KERATIN.
 DR PROSITE; PS00226; IF; UNKNOWN_1.
 KW Intermediate filament.
 SQ SEQUENCE 525 AA; 55145 MW; A97E16C09045B928 CRC64;

 Query Match 97.5%; Score 39; DB 2; Length 525;
 Best Local Similarity 85.7%; Pred. No. 3.4e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

 QY 1 GGIGDGG 7
 DB 128 GGVGDGG 134

 RESULT 27
 Q8LNH8 ORYSA
 ID Q8LNH8 ORYSA PRELIMINARY; PRT; 1291 AA.
 AC Q8LNH8;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
 DE Putative glycine-rich protein.
 GN ORFNames=OSJNBa0078001.2;
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Kim M.M.,
 RA Overton II L.L., Bera J.J., Tsitrin T., Krol M.I., Jarrahi B.B.,
 RA Jin S.S., Koo H., Zismann V., Hsiao J., Blunt S., Vanaken S.S.,
 RA Utterback T.T., Feldblyum T.V., Yang Q.Q., Haas B.J., Suh B.B.,
 RA Peterson J.J., Quackenbush J., White O., Salzberg S.L., Fraser C.M.;
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Buell R.;
 RA Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RA The Rice Chromosome 10 Sequencing Consortium;
 RT "In-depth view of structure, activity, and evolution of rice
 RT chromosome 10.";
 RL Science 300:1566-1569(2003).
 RN [4]
 RP NUCLEOTIDE SEQUENCE.
 RA Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.;
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC079888; AAM93675.1; -; Genomic DNA.
 DR EMBL; AB017109; AAP54457.1; -; Genomic DNA.
 DR Gramene; Q8LNH8; -.
 SQ SEQUENCE 1291 AA; 137489 MW; C28CEBED452C529A CRC64;

 Query Match 97.5%; Score 39; DB 2; Length 1291;
 Best Local Similarity 85.7%; Pred. No. 8.5e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

 QY 1 GGIGDGG 7
 DB 54 GGVGDGG 60

 RESULT 28
 Q7SBF8 NEUCR
 ID Q7SBF8 NEUCR PRELIMINARY; PRT; 1356 AA.
 AC Q7SBF8;
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Hypothetical protein.
GN Name=NCU07598.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=OR74A;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., Fitzhugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Iankiev P., Pedersen D., Nelson M., Washburne M.,
RA Seitzrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
RA Kothé G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
RA Kamal M., Kamvysselis M., Mauceli E., Bielke C., Rudd S., Frishman D.,
RA Krystofova S., Rasmussen C., Metznerberg R.L., Perkins D.D., Kroken S.,
RA Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Omani S.A.,
RA DeSouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
RA Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.,
RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
RL Nature 0:0-0(2003).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -!- SIMILARITY: Belongs to the class-V pyridoxal-phosphate-dependent
CC aminotransferase family.
CC EMBL; ABX01000168; EAA3742.1; -; Genomic DNA.
DR GO; GO:0015036; F:disulfide oxidoreductase activity; IEA.
DR GO; GO:0008483; F:transaminase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR000192; AminoTrans_V.
DR InterPro; IPR002938; mOase_FAD_Bnd.
DR InterPro; IPR003042; Rng_hydrolase.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF00266; AminoTrans_5; 1.
DR Pfam; PF01494; FAD_binding_3; 1.
DR PRINTS; PR00420; RNMNOMGNASE.
DR PRINTS; PR00453; VWFADOMAIN.
KW Hypothetical protein.
SQ SEQUENCE 1356 AA; 151112 MW; 940ESA38D62C5FD9 CRC64;
Query Match 97.5%; Score 39; DB 2; Length 1356;
Best Local Similarity 85.7%; Pred. No. 8.9e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGIGDGG 7
DB 792 GGVGGG 798
||:|||||
RESULT 29
Q9LGG6 ORYSA
ID Q9LGG6 ORYSA PRELIMINARY; PRT; 214 AA.
AC Q9LGG6;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-FEB-2005 (TREMBlrel. 29, Last annotation update)
DE Hypothetical protein (OSJNBa0036E02.19 protein).
GN Name=OSJNBa0036E02.19;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.

RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP002804; BAB00649.1; -; Genomic DNA.
DR EMBL; AP002862; BAB17745.1; -; Genomic DNA.
DR Gramene; Q9LGG6; -;
KW Hypothetical protein.
SQ SEQUENCE 214 AA; 22860 MW; E5DABE7661E453D5 CRC64;
Query Match 95.0%; Score 38; DB 2; Length 214;
Best Local Similarity 85.7%; Pred. No. 2.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGIGDGG 7
DB 98 GGLGDGG 104
||:|||||
RESULT 30
Q5BIA5 DROME
ID Q5BIA5 DROME PRELIMINARY; PRT; 228 AA.
AC Q5BIA5;
DT 10-MAY-2005 (TREMBlrel. 30, Created)
DT 10-MAY-2005 (TREMBlrel. 30, Last sequence update)
DE RE09303p.
GN Name=beat-1a;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Berkley;
RA Stapleton M., Carlson J., Chavez C., Frise E., George R., Pacleb J.,
RA Park S., Wan K., Yu C., Rubin G.M., Celniker S.;
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; BT021319; AAX33467.1; -; mRNA.
DR InterPro; IPR011992; EF-Hand type.
SQ SEQUENCE 228 AA; 25221 MW; E48D64C535B8D20B CRC64;
Query Match 95.0%; Score 38; DB 2; Length 228;
Best Local Similarity 85.7%; Pred. No. 2.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGIGDGG 7
DB 97 GGLGDGG 103
||:|||||
RESULT 31
Q75I35 ORYSA
ID Q75I35 ORYSA PRELIMINARY; PRT; 273 AA.
AC Q75I35;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 01-FEB-2005 (TREMBlrel. 29, Last annotation update)
DE Hypothetical protein Os03g40910 (Hypothetical protein
DE OSJNBa0004G03.26).
GN Name=Os03g40910; Synonyms=OSJNBa0004G03.26;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA Overton II L.B., Tsitrin T., Kim M.M., Bera J.J., Jin S.,
RA Fadrosch D.W., Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S.,
RA Vanaken S.S., Riedmuller S.B., Utterback T.T., Feidblyum T.V.,

```

RA Yang Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J., White O.,
RA Salzberg S.L., Frazer C.M.;
RT "Oryza sativa chromosome 3 BAC OSJNBa0091E13 genomic sequence.";
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
[2]
RP NUCLEOTIDE SEQUENCE.
RA Buell R.;
RL Submitted (JAN-2005) to the EMBL/GenBank/DBJ databases.
[3]
RN NUCLEOTIDE SEQUENCE.
RP Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA Overton II L.L., Teitlin T., Kim M.M., Bera J.J., Jin S.S.,
RA Padroon D.W., Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S.,
RA Vanaken S.S., Riedmuller S.B., Utterback T.T., Feldblyum T.V.,
RA Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
RA White O., Salzberg S.L., Frazer C.M.;
RT "Oryza sativa chromosome 3 BAC OSJNBa004G03 genomic sequence.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
[4]
RN NUCLEOTIDE SEQUENCE.
RP Buell R.;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC133860; AAR87235.1; -; Genomic DNA.
DR EMBL; AC109601; AAT78765.1; -; Genomic DNA.
DR Gramene; Q75135; -; NmrA.
DR InterPro; IPR008030; NmrA.
DR Pfam; PF05368; NmrA; 1.
KW Hypothetical protein.
SQ SEQUENCE 273 AA; 27222 MW; 7D4FC277A0B86DD7 CRC64;

Query Match 95.0%; Score 38; DB 2; Length 273;
Best Local Similarity 85.7%; Pred. No. 2.6e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGLGDGG 7
DB 81 GGLGDGG 87

RESULT 32
Q6YU4_ORYSA
ID Q6YU4_ORYSA PRELIMINARY; PRT; 300 AA.
AC Q6YU4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Hypothetical protein OJ1134_E08.29.
GN Name=OJ1134_E08.29;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Hattori M., Sakaki Y., Katayose Y.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 9, BAC
clone OJ1134_E08."
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005860; BAD16501.1; -; Genomic DNA.
DR Gramene; Q6YU4; -;
KW Hypothetical protein.
SQ SEQUENCE 300 AA; 32403 MW; 123CA3AFA7FAP8 CRC64;

Query Match 95.0%; Score 38; DB 2; Length 300;
Best Local Similarity 85.7%; Pred. No. 2.9e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGLGDGG 7
DB 188 GGLGDGG 194

```

```

RESULT 33
Q94534_DROME
ID Q94534_DROME PRELIMINARY; PRT; 427 AA.
AC Q94534;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Beaten path precursor (beat protein) (CG4846-PA).
GN Name=beat-1a; Synonyms=beat; ORFNames=CG4846;
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Fambrough D., Goodman C.S.;
RT "The Drosophila beaten path Gene Encodes A Novel Secreted Protein That
Regulates Defasciculation At Motor Axon Choice Points.";
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
[2]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=Berkley;
RX MEDLINE=99403001; PubMed=10471707;
RA Ashburner M., Misra S., Roote J., Lewis S.E., Blazej R.G., Davis T.,
RA Doyle C., Galle R.F., George R.A., Harris N.L., Hartzell G.,
RA Harvey D.A., Hong L., Houston K.A., Hoskins R.A., Johnson G.,
RA Martin C., Moshrefi A.R., Palazzolo M., Reese M.G., Spradling A.C.,
RA Tsang G., Wan K.H., Whitelaw K., Celniker S.E., Rubin G.M.;
RT "An exploration of the sequence of a 2.9-Mb region of the genome of
Drosophila melanogaster: the Adh region.";
RL Genetics 153:179-219 (1999).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RP STRAIN=Berkley;
RA Celniker S.E., Agbayani A., Arcaina T.T., Baxter E., Blazej R.G.,
RA Butenhoff C., Chavez M., Chew M., Ciesiolka L., Doyle C.M.,
RA Farfan D.E., Galle R., George R.A., Harris N.L., Hoskins R.A.,
RA Houston K.A., Hummasti S.R., Karra K., Kearney L., Kim E., Lee B.,
RA Lewis S., Li P., Lomotan M.A., Mazda P., Moshrefi A.R., Moshrefi M.,
RA Nixon K., Pacleb J.M., Park S., Pfeiffer B., Poon L., Sequeira A.,
RA Sethi H., Snir E., Svirskas R.R., Wan K.H., Weinburg T., Zhang R.,
RA Zieran L., Rubin G.M.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[4]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=Berkley;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brannon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadiou E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
RA Fostel C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Hewland T.J., Wei M.-H., Ibegwam C.,
RA Hostin D., Houston K.A., Howland T.J., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kravitz S., Kulp D., Lai Z.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,

```

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Waasman D.A., Weinstein G.M., Weisenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RT Science 287:2185-2195(2000).
RN [5]
RN NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: release 3 of the *Drosophila*
RT melanogaster euchromatic genome sequence.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [6]
RN NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the *Drosophila melanogaster* euchromatin:
RT a genomic perspective.";
RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
RN [7]
RN NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bertram B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN [8]
RN NUCLEOTIDE SEQUENCE.
RG Berkeley *Drosophila* Genome Project;
RA Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
RA Hoskins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,
RA Yu C., Rubin G.;
RT "*Drosophila melanogaster* release 4 sequence.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN [9]
RN NUCLEOTIDE SEQUENCE.
RG FlyBase;
RL Submitted (MAR-2005) to the EMBL/GenBank/DDBJ databases.
CC -!- INTERACTION:
CC Q9VFC8:CG6904; NbExp=1; IntAct=EBI-113385; EBI-169205;
DR EMBL; U67057; AAB07545.1; -, mRNA.
DR EMBL; AE003415; AAF44982.1; -, Genomic DNA.
DR EMBL; AE003649; AAF53500.1; -, Genomic DNA.
DR InAct; Q94534; -,
DR FlyBase; FBgn0013433; beat-1a.
DR GO; GO:0016199; P:axon choice point recognition; TAS.
DR GO; GO:0001746; P:Boltwig's organ morphogenesis; IMP.
DR GO; GO:0007415; P:defasciculation of motor neuron; IMP.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR SMART; SM00409; IG; 1.
DR PROSITE; PS50835; IG LIKE; 1.
DR Immunoglobulin domain; Signal.

FT SIGNAL 1 26 Potential.
FT CHAIN 27 427 beaten path.
SQ SEQUENCE 427 AA; 47917 MW; 26646C37D88B119E CRC64;
Query Match 95.0%; Score 38; DB 2; Length 427;
Best Local Similarity 85.7%; Pred. No. 4.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGIGDGG 7
DB 296 GGLGDGG 302
RESULT 34
Q4H937_9DEIO PRELIMINARY; PRT; 664 AA.
ID Q4H937_9DEIO PRELIMINARY; PRT; 664 AA.
AC Q4H937_9DEIO PRELIMINARY; PRT; 664 AA.
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein precursor.
GN ORFNames=bgodRAFT 1068;
OS *Deinococcus geothermalis* DSM 11300.
OC Bacteria; *Deinococcus-Thermus*; *Deinococci*; *Deinococcales*;
OC *Deinococcaceae*; *Deinococcus*.
OX NCBI_TaxID=319795;
RN [1]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=DSM 11300;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
RA Hammon N., Izrati S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of *Deinococcus geothermalis*
RT DSM 11300.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.
RN [2]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=DSM 11300;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of *Deinococcus geothermalis*
RT DSM 11300.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DDBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAHE01000005; EAL82815.1; -, Genomic DNA.
KW Hypothetical protein; Signal.
FT SIGNAL 454 476 Potential.
SQ SEQUENCE 664 AA; 71078 MW; CAF4C31CE20B9464 CRC64;
Query Match 95.0%; Score 38; DB 2; Length 664;
Best Local Similarity 85.7%; Pred. No. 6.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGIGDGG 7
DB 146 GGLGDGG 152
RESULT 35
Q4H3G4_CIOIN PRELIMINARY; PRT; 200 AA.
ID Q4H3G4_CIOIN PRELIMINARY; PRT; 200 AA.
AC Q4H3G4_CIOIN PRELIMINARY; PRT; 200 AA.
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Ci-FUSE protein.
GN Name=Ci-FUSE;
OS *Ciona intestinalis*.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
OC *Phlebobranchia*; *Cionidae*; *Ciona*.
OX NCBI_TaxID=7719;

RN NUCLEOTIDE SEQUENCE.
RX PubMed=15269171; DOI=10.1242/dev.01270; Satou Y.;
RA Imai K.S., Hino K., Yagi K., Satoh N., Satou Y.;
RT "Genomewide surveys of developmentally relevant genes in Ciona
intestinalis.";
RL Development 131:4047-4058(2004).
(2)
RN NUCLEOTIDE SEQUENCE.
RX PubMed=12736827; DOI=10.1007/s00427-003-0330-z;
RA Satou Y., Satoh N.;
RT "Genomewide surveys of developmentally relevant genes in Ciona
intestinalis.";
RL Dev. Genes Evol. 213:211-212(2003).
(3)
RN NUCLEOTIDE SEQUENCE.
RA Satou Y.;
RT "Expressed genes in Ciona intestinalis.";
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB210458; BAE06462.1; -; mRNA.
SQ SEQUENCE 200 AA; 21821 MW; DE42E2BA12E2C8E4 CRC64;

Query Match 92.5%; Score 37; DB 2; Length 200;
Best Local Similarity 85.7%; Pred. No. 2.8e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
||:||||
Db 127 GGMGDGG 133

RESULT 36
Q4H3G5 CIOIN
ID Q4H3G5 CIOIN PRELIMINARY; PRT; 325 AA.
AC Q4H3G5;
DT 13-SEP-2005 (T-EMBLrel. 31, Created)
DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)
DE Ci-FUSE protein.
GN Name=Ci-FUSE;
OS Ciona intestinalis.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
OC Phlebobranchia; Cionidae; Ciona.
OX NCBI_TaxID=7719;
(1)
RN NUCLEOTIDE SEQUENCE.
RX PubMed=15269171; DOI=10.1242/dev.01270; Satou Y.;
RA Imai K.S., Hino K., Yagi K., Satoh N., Satou Y.;
RT "Genomewide surveys of developmentally relevant genes in Ciona
intestinalis.";
RL Development 131:4047-4058(2004).
(2)
RN NUCLEOTIDE SEQUENCE.
RX PubMed=12736827; DOI=10.1007/s00427-003-0330-z;
RA Satou Y., Satoh N.;
RT "Genomewide surveys of developmentally relevant genes in Ciona
intestinalis.";
RL Dev. Genes Evol. 213:211-212(2003).
(3)
RN NUCLEOTIDE SEQUENCE.
RA Satou Y.;
RT "Expressed genes in Ciona intestinalis.";
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB210457; BAE06462.1; -; mRNA.
SQ SEQUENCE 325 AA; 34691 MW; C679610AC3FBA2EB CRC64;

Query Match 92.5%; Score 37; DB 2; Length 325;
Best Local Similarity 85.7%; Pred. No. 4.6e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
||:||||
Db 135 GGMGDGG 141

RESULT 37
Q51PIL MAGGR
ID Q51PIL MAGGR PRELIMINARY; PRT; 405 AA.
AC Q51PIL;
DT 13-SEP-2005 (T-EMBLrel. 31, Created)
DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (T-EMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=MG00495.4;
OS Magnaporthe grisea 70-15.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.
OX NCBI_TaxID=242507;
(1)
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA Ait-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Ambruster J., Bachantsang P., Baldwin J., Barry A.,
RA Bayul T., Blitshteyn B., Bloom I., Blye J., Boguslavskiy L.,
RA Borowsky M., Boukhgalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,
RA Gnrke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Kamat A., Kamyssealis M., Karlsson E.,
RA Kells C., Kieu A., Kisner P., Kodira C., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O.,
RA Lui A., Ma L.-J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,
RA McCarthy M., McDonough S., Mcghee T., Meldrum J., Meneus L.,
RA Mesirov J., Mihalev A., Mihova T., Mikkelsen T., Mlenga V., Moru K.,
RA Mozes J., Mulrain L., Munson G., Naylor J., News C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okoawo O., O'leary S., Omotosho B.,
RA O'Neill K., Osman S., Parker S., Perrin D., Phunkhang P., Pigani B.,
RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schubach R., Seaman C., Settipalli S., Sharpe T.,
RA Sheridan B., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
RA Stetson K., Stalker J., Stange-thomann N., Stavropoulos S.,
RA Tenzing P., Tesfaye S., Stone C., Stubbs M., Talamas J., Tchuanga P.,
RA Towey S., Tsamta T., Tsomo N., Vallee D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
RA Zimmer A., Zody M., Lander E.;
RT "The genome sequence of Magnaporthe grisea.";
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
(2)
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Dean R., Mitchell T., Brown D., Pan H., Thon M.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
(3)
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Zhu H., Blackmon B.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACU01001423; EAA48837.1; -; Genomic_DNA.
KW Hypothetical protein.

SQ SEQUENCE 405 AA; 43344 MW; 399A66BA8DBE1308 CRC64;

Query Match 92.5%; Score 37; DB 2; Length 405;
Best Local Similarity 85.7%; Pred. No. 5.7e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
||:||||
Db 319 GGMGDGG 325

RESULT 38

Q7UXR4 RHOB4 PRELIMINARY; PRT; 417 AA.
AC Q7UXR4;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-WAR-2004 (TREMBlrel. 26, Last annotation update)
DE Pleiotropic regulatory protein.
GN Name=degri; OrderedLocustNames=RB1170;
OS Rhodopirellula baltica.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1;
RX MEDLINE=2735913; PubMed=12835416; DOI=10.1073/pnas.1431443100;
RA Gieseckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heilmann K., Rabus R.,
RA Schlesner H., Amann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
RT strain 1.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
DR EMBL; BX294134; CAD71939.1; -; Genomic_DNA.
DR InterPro; IPR000653; DegT_DnrJ_EryCl.
DR Pfam; PF01041; DegT_DnrJ_EryCl; 1.
KW Complete proteome.
SQ SEQUENCE 417 AA; 45684 MW; 315531D637D8108E CRC64;

Query Match 92.5%; Score 37; DB 2; Length 417;
Best Local Similarity 85.7%; Pred. No. 5.9e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
||:||||
Db 225 GGMGDGG 231

RESULT 39

Q9DUC2_9VIRU PRELIMINARY; PRT; 100 AA.
AC Q9DUC2;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 23, Last annotation update)
DE ORF2.
GN Name=ORF2;
OS Torque teno virus.
OC Viruses; ssDNA viruses; Anellovirus.
OX NCBI_TaxID=68887;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20534983; PubMed=11080484; DOI=10.1006/viro.2000.0588;
RA Okamoto H., Nishizawa T., Tawara A., Peng Y., Takahashi M.,
RA Kishimoto J., Tanaka T., Miyakawa Y., Mayumi M.;
RT "Species-specific TT viruses in humans and nonhuman primates and their
RT phylogenetic relatedness.";
RL Virology 277:368-378(2000).
DR EMBL; AB041960; BAB19315.1; -; Genomic_DNA.
DR InterPro; IPR004118; TT_ORF2.
DR Pfam; PF02957; TT_ORF2; 1.
SQ SEQUENCE 100 AA; 10424 MW; 514E4DF088112308 CRC64;

Query Match 90.0%; Score 36; DB 2; Length 100;
Best Local Similarity 85.7%; Pred. No. 2.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
|||:|
Db 50 GGIGEGG 56

RESULT 40

Q5Z747 ORYSA PRELIMINARY; PRT; 130 AA.
AC Q5Z747;
DT 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Hypothetical protein OSJNBa0012F14.33.
GN Name=OSJNBa0012F14.33;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, BAC
RT clone:OSJNBa0012F14.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP004784; BAD61892.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 130 AA; 11675 MW; AF98D36005C747BB CRC64;

Query Match 90.0%; Score 36; DB 2; Length 130;
Best Local Similarity 85.7%; Pred. No. 2.7e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
|||:|
Db 40 GGIGEGG 46

RESULT 41

Q9C6Y9_ARATH PRELIMINARY; PRT; 223 AA.
AC Q9C6Y9;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-FEB-2005 (TREMBlrel. 29, Last annotation update)
DE Hypothetical protein T7023.13 (Hypothetical protein T18F15.12).
GN Name=T7023.13; Synonyms=T18F15.12;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
RA Maiti R., Ronning C.M., Koo H., Fujii C.Y., Uterback T.R.,
RA Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
RA Maiti R., Ronning C.M., Koo H., Fujii C.Y., Uterback T.R.,
RA Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC074228; AAG50547.1; -; Genomic_DNA.
DR EMBL; AC084807; AAK43482.1; -; Genomic_DNA.
DR PIR; B96506; B96506.
DR InterPro; IPR002952; Eggshell.

```

DR PRINTS; PRO1228; EGGSHLL.
KW Hypothetical protein.
SQ SEQUENCE 223 AA; 19781 MW; E21A0F404896897A CRC64;

  Query Match          90.0%; Score 36; DB 2; Length 223;
  Best Local Similarity 85.7%; Pred. No. 4.6e+02;
  Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
Db 212 GGFEGGG 218

RESULT 42
Q9A971 CAUCR
ID Q9A971 CAUCR PRELIMINARY; PRT; 224 AA.
AC Q9A971; 2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein CC1123.
GN OrderedLocusNames=CC1123;
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
OC Caulobacteraceae; Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647; DOI=10.1073/pnas.061029298;
RA Nieman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J.A., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H.M., Shetty J.,
RA Berry K.J., Utterback T.R., Tran K., Wolf A.M., Vamathevan J.J.,
RA Ermolaeva M.D., White O., Salzberg S.L., Venter J.C., Shapiro L.,
RA Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL; AS005790; AAK23107.1; -; Genomic_DNA.
DR PIR; G87388; G87388.
DR TIGR; CC1123; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 224 AA; 23458 MW; FCC2CE335859216C CRC64;

  Query Match          90.0%; Score 36; DB 2; Length 224;
  Best Local Similarity 85.7%; Pred. No. 4.6e+02;
  Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
Db 149 GGFEGGG 155

RESULT 43
Q4P316 USTMA
ID Q4P316 USTMA PRELIMINARY; PRT; 303 AA.
AC Q4P316; 2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=UM05497.1;
OS Ustilago maydis 521.
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
OX NCBI_TaxID=237631;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=521;
RA Birren B., Nusbaum C., Abebe A., Abouelheil A., Adekoya E.,
RA Ait-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,

```

```

RA Arachchi H., Armbruster J., Bachantsang P., Baldwin J., Barry A.,
RA Bayul T., Blitshsteyn B., Bloom T., Blye J., Boguslavskiy L.,
RA Borowsky M., Boukhgalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA Collimore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,
RA Gnirke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Kamat A., Kamvysseis M., Karlsson E.,
RA Kellis C., Kieu A., Kisner P., Kodira C., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,
RA McCarthy M., McDonough S., Mcghee T., Meldrum J., Meneus L.,
RA Mesirov J., Mihalov A., Mihova T., Mikkelsen T., Mlenga V., Moru K.,
RA Mozes J., Mulrain L., Munson G., Raylor J., News C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okoawo O., O'leary S., Omotosho B.,
RA O'Neill K., Osman S., Parker S., Perrin D., Phunkhang P., Pigani B.,
RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Rogov P.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schnupbach R., Seaman C., Settippalli S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA Tenzing P., Tesfaye S., Theodore J., Thoulutsang Y., Topham K.,
RA Towey S., Tsamla T., Tsomo N., Vallee D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
RA Zimmer A., Zody M., Lander E.;
RT "The genome sequence of Ustilago maydis."
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACP01000197; EAK86430.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 303 AA; 31887 MW; CC14044903AA9BA6 CRC64;

  Query Match          90.0%; Score 36; DB 2; Length 303;
  Best Local Similarity 85.7%; Pred. No. 6.3e+02;
  Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
Db 267 GGIGEGG 273

RESULT 44
O68842 STRAT
ID O68842 STRAT PRELIMINARY; PRT; 368 AA.
AC O68842;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Amino transferase.
GN Name=oleN2;
OS Streptomyces antibioticus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1890;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 11891;
RX MEDLINE=98343801; PubMed=9680207;
RA Quiros L.M., Aguirrezabalaga I., Olano C., Mendez C., Salas J.A.;
RT "Two glycosyltransferases and a glycosidase are involved in

```

```

RT oleandomycin modification during its biosynthesis by Streptomycetes
RT antibioticus".; 28:1177-1185(1998).
RL Mol. Microbiol. 28:1177-1185(1998).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 11891.
RA Quiros L.M., Aguirrezabalaga I., Olano C., Mendez C., Salas J.A.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF055579; AAD55458.1; -; Genomic_DNA.
DR F01; T51111; T51111.
DR HSP; Q82NF3; 1MD0.
DR GO; GO:0008483; F:transaminase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR000653; Degt_DnrJ_EryC1.
DR Pfam; PF01041; Degt_DnrJ_EryC1; 1.
KW Aminotransferase; Transferase.
SQ SEQUENCE 368 AA; 39601 MW; 7911658830FA0C3F CRC64;

Query Match 90.0%; Score 36; DB 2; Length 368;
Best Local Similarity 85.7%; Pred. No. 7.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
DB 188 GGFGDGG 194

RESULT 45
Q9L6B9 STRAT
ID Q9L6B9_STRAT PRELIMINARY; PRT; 368 AA.
AC Q9L6B9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Aminotransferase-like protein.
OS Streptomycetes antibioticus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=1890;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Tu99;
RA Draeger G., Park S.H., Floss H.G.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF237895; AAF59939.1; -; Genomic_DNA.
DR HSP; Q82NF3; 1MD0.
DR GO; GO:0008483; F:transaminase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR000653; Degt_DnrJ_EryC1.
DR Pfam; PF01041; Degt_DnrJ_EryC1; 1.
KW Aminotransferase; Transferase.
SQ SEQUENCE 368 AA; 39526 MW; 7911759201D2153E CRC64;

Query Match 90.0%; Score 36; DB 2; Length 368;
Best Local Similarity 85.7%; Pred. No. 7.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
DB 188 GGFGDGG 194

RESULT 46
Q6L561 ORYSA
ID Q6L561_ORYSA PRELIMINARY; PRT; 383 AA.
AC Q6L561;

```

```

DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Putative nuclear RNA binding protein A (Hypothetical protein
DE OSJNBa0022J22.4)
GN Name=OJ1115_D04.10; Synonyms=OSJNBa0022J22.4;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartioideae; Oryzaeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Chow T.-Y., Hsing Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M.,
RA Chao Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,
RA Chen Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,
RA Hsiung J.-N., Hsu C.-H., Huang J.-J., Kau P.-I., Lee M.-C.,
RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,
RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,
RA Wu H.-P., Shaw J.-F.;
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Chow T.-Y., Hsing Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M.,
RA Chao Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,
RA Chen Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,
RA Hsiung J.-N., Hsu C.-H., Huang J.-J., Kau P.-I., Lee M.-C.,
RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,
RA Wu H.-P., Shaw J.-F.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC105260; AAT38003.1; -; Genomic_DNA.
DR EMBL; AC136216; AAT58811.1; -; Genomic_DNA.
DR Gramene; Q6L561; -.
DR InterPro; IPR006861; HARP4_PAI-RBP1.
DR Pfam; PF04774; HARP4_PAI-RBP1; 1.
KW Hypothetical protein.
SQ SEQUENCE 383 AA; 41029 MW; 83DAECAD42C66810 CRC64;

Query Match 90.0%; Score 36; DB 2; Length 383;
Best Local Similarity 85.7%; Pred. No. 7.9e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGIGDGG 7
DB 112 GGFGDGG 118

RESULT 47
Q69SQ7 ORYSA
ID Q69SQ7_ORYSA PRELIMINARY; PRT; 405 AA.
AC Q69SQ7;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein OSJNBa0016019.23.
GN Name=OSJNBa0016019.23;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartioideae; Oryzaeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP004991; BAD35978.1; -; Genomic_DNA.
DR Gramene; Q69SQ7; -.
DR InterPro; IPR000408; Reg chr condens.
DR PROSITE; PS00626; RCCL_2; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 405 AA; 41286 MW; 287B9052677EB3AF CRC64;

```

```

Query Match          90.0%; Score 36; DB 2; Length 405;
Best Local Similarity 85.7%; Pred. No. 8.4e+02;
Matches      6; Conservative      1; Mismatches      0; Indels      0; Gaps      0;

QY      1 GGIGDGG 7
DB      66 GGIGEGG 72
      1  |||||:|
      66  |||||:|

RESULT 48
Q5RJ37_BRARE
ID      Q5RJ37_BRARE PRELIMINARY;      PRT;      408 AA.
AC      Q5RJ37;
DT      01-FEB-2005 (TrEMBLrel. 29, Created)
DT      01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT      01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE      Novel protein similar to vertebrate glutamic pyruvate transaminase
DE      (Alanine aminotransferase) 2 (GPT2) (Fragment).
GN      NAME=OTDARP0000006514; ORFNames=CH211-57117.3-001;
OS      Brachydanio rerio (zebrafish) (Danio rerio).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC      Cyprinidae; Danio.
OX      NCBI_TaxID=7955;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RA      Hall R.;
RL      Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
DR      EMBL; BX004824; CAIL2022.1; -; Genomic DNA.
DR      GO; GO:0008483; F:transaminase activity; IEA.
DR      GO; GO:0016740; F:transferase activity; IEA.
KW      Aminotransferase; Pyruvate; Transferase.
FT      NON TER      408      408
FT      SEQUENCE      408 AA;      44281 MW;      4B126330F442E715 CRC64;

Query Match          90.0%; Score 36; DB 2; Length 408;
Best Local Similarity 85.7%; Pred. No. 8.4e+02;
Matches      6; Conservative      0; Mismatches      1; Indels      0; Gaps      0;

QY      1 GGIGDGG 7
DB      21 GGFGDGG 27
      1  |||||
      21  GGFGDGG 27

RESULT 49
Q4ZLL5_PSESY
ID      Q4ZLL5_PSESY PRELIMINARY;      PRT;      424 AA.
AC      Q4ZLL5;
DT      13-SEP-2005 (TrEMBLrel. 31, Created)
DT      13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT      13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE      Outer membrane porin.
OR      ORFNames=Psy_4930;
OS      Pseudomonas syringae pv. syringae B728a.
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC      Pseudomonadaceae; Pseudomonas.
OX      NCBI_TaxID=205918;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=B728a;
RG      DOE Joint Genome Institute;
RA      Chain P.; Latimer F.; DiBartolo G.; Copeland A.; Lykidis A.; Trong S.;
RA      Nolan M.; Goltsman E.; Thiel J.; Malfatti S.; Lapidus A.; Detter J.C.;
RA      Land M.; Richardson P.M.; Kyrides N.C.; Ivanova N.;
RT      "Comparison of two complete genome sequences of Pseudomonas syringae
RT      pv. syringae B728a and pv. tomato DC3000."
RL      Proc. Natl. Acad. Sci. U.S.A. 0:0-0(2005).
RN      [2]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=B728a;
RG      Loper J.;
RA      Loper J.;
RL      Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN      [3]

```